VOLUME 5,NUMBER 1 JANUARY 2007

ISSN:1548-5390 PRINT,1559-176X ONLINE

JOURNAL

OF CONCRETE

AND APPLICABLE

MATHEMATICS

EUDOXUS PRESS,LLC

SCOPE AND PRICES OF THE JOURNAL
Journal of Concrete and Applicable Mathematics

A quartely international publication of **Eudoxus Press,LLC**

Editor in Chief: George Anastassiou
Department of Mathematical Sciences,
University of Memphis
Memphis, TN 38152, U.S.A.
ganastss@memphis.edu

The main purpose of the "Journal of Concrete and Applicable Mathematics" is to publish high quality original research articles from all subareas of Non-Pure and/or Applicable Mathematics and its many real life applications, as well connections to other areas of Mathematical Sciences, as long as they are presented in a Concrete way. It welcomes also related research survey articles and book reviews.A sample list of connected mathematical areas with this publication includes and is not restricted to: Applied Analysis, Applied Functional Analysis, Probability theory, Stochastic Processes, Approximation Theory, O.D.E, P.D.E, Wavelet, Neural Networks,Difference Equations, Summability, Fractals, Special Functions, Splines, Asymptotic Analysis, Fractional Analysis, Inequalities, Moment Theory, Numerical Functional Analysis,Tomography, Asymptotic Expansions, Fourier Analysis, Applied Harmonic Analysis, Integral Equations, Signal Analysis, Numerical Analysis, Optimization, Operations Research, Linear Programming, Fuzzyness, Mathematical Finance, Stochastic Analysis, Game Theory, Math.Physics aspects, Applied Real and Complex Analysis, Computational Number Theory, Graph Theory, Combinatorics,
Computer Science Math.related topics,combinations of the above, etc. In general any kind of Concretely presented Mathematics which is Applicable fits to the scope of this journal.
Working Concretely and in Applicable Mathematics has become a main trend in many recent years,so we can understand better and deeper and solve the important problems of our real and scientific world.
"Journal of Concrete and Applicable Mathematics" is a peer- reviewed International Quarterly Journal.
We are calling for papers for possible publication. The contributor should send three copies of the contribution to the editor in-Chief typed in TEX, LATEX double spaced. [See: Instructions to Contributors]

Journal of Concrete and Applicable Mathematics(JCAAM)
ISSN:1548-5390 PRINT, 1559-176X ONLINE.
is published in January,April,July and October of each year by
EUDOXUS PRESS,LLC,
1424 Beaver Trail Drive,Cordova,TN38016,USA,
Tel.001-901-751-3553
anastassioug@yahoo.com
http://www.EudoxusPress.com.
Visit also www.msci.memphis.edu/~ganastss/jcaam.
Webmaster:Ray Clapsadle

Annual Subscription Current Prices:For USA and Canada,Institutional:Print $250,Electronic $220,Print and Electronic $310.Individual:Print $77,Electronic $60,Print

&Electronic $110.For any other part of the world add $25 more to the above prices for Print.
Single article PDF file for individual $8.Single issue in PDF form for individual $25.
The journal carries page charges $10 per page of the pdf file of an article,payable upon acceptance of the article within one month and before publication.
No credit card payments.Only certified check,money order or international check in US dollars are acceptable.
Combination orders of any two from JoCAAA,JCAAM,JAFA receive 25% discount,all three receive 30% discount.

JCAAM is reviewed and abstracted by AMS Mathematical Reviews,MATHSCI,and Zentralblaat MATH.

JCAAM IS A JOURNAL OF RAPID PUBLICATION

9) A.M.Fink
Department of Mathematics
Iowa State University
Ames,IA 50011-0001,USA
tel.515-294-8150
fink@math.iastate.edu
Inequalities,Ordinary Differential
Equations

10) Sorin Gal
Department of Mathematics
University of Oradea
Str.Armatei Romane 5
3700 Oradea,Romania
galso@uoradea.ro
Approximation Th.,Fuzzyness,Complex
Analysis

11) Jerome A.Goldstein
Department of Mathematical Sciences
The University of Memphis,
Memphis,TN 38152,USA
tel.901-678-2484
jgoldste@memphis.edu
Partial Differential Equations,
Semigroups of Operators

12) Heiner H.Gonska
Department of Mathematics
University of Duisburg
Duisburg,D-47048
Germany
tel.0049-203-379-3542 office
gonska@informatik.uni-duisburg.de
Approximation Th.,Computer Aided
Geometric Design

13) Dmitry Khavinson
Department of Mathematical Sciences
University of Arkansas
Fayetteville,AR 72701,USA
tel.(479)575-6331,fax(479)575-8630
dmitry@uark.edu
Potential Th.,Complex Analysis,Holomorphic
PDE,
Approximation Th.,Function Th.

14) Virginia S.Kiryakova
Institute of Mathematics and Informatics
Bulgarian Academy of Sciences
Sofia 1090,Bulgaria
virginia@diogenes.bg
Special Functions,Integral Transforms,
Fractional Calculus

15) Hans-Bernd Knoop

29) Bl. Sendov
Institute of Mathematics and Informatics
Bulgarian Academy of Sciences
Sofia 1090,Bulgaria
bsendov@bas.bg
Approximation Th.,Geometry of Polynomials,
Image Compression

30) Igor Shevchuk
Faculty of Mathematics and Mechanics
National Taras Shevchenko
University of Kyiv
252017 Kyiv
UKRAINE
shevchuk@univ.kiev.ua
Approximation Theory

31) H.M.Srivastava
Department of Mathematics and Statistics
University of Victoria
Victoria,British Columbia V8W 3P4
Canada
tel.250-721-7455 office,250-477-6960 home,
fax 250-721-8962
harimsri@math.uvic.ca
Real and Complex Analysis,Fractional Calculus
and Appl.,
Integral Equations and Transforms,Higher
Transcendental
Functions and Appl.,q-Series and q-Polynomials,
Analytic Number Th.

32) Ferenc Szidarovszky
Dept.Systems and Industrial Engineering
The University of Arizona
Engineering Building,111
PO.Box 210020
Tucson,AZ 85721-0020,USA
szidar@sie.arizona.edu
Numerical Methods,Game Th.,Dynamic Systems,
Multicriteria Decision making,
Conflict Resolution,Applications
in Economics and Natural Resources
Management

33) Gancho Tachev
Dept.of Mathematics
Univ.of Architecture,Civil Eng. and Geodesy
1 Hr.Smirnenski blvd
BG-1421 Sofia,Bulgaria
gtt_fte@uacg.bg
Approximation Theory

34) Manfred Tasche
Department of Mathematics

JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS,VOL.5,NO.1,9-24,2007,COPYRIGHT 2007 EUDOXUS PRESS,LLC

Extensions of Montgomery identity with applications for α-L-Hölder type functions

A. Aglić Aljinović and J. Pečarić

Department of applied mathematics
Faculty of Electrical Engineering and Computing
Unska 3, 10 000 Zagreb
Croatia
andrea@zpm.fer.hr

Faculty of textile technology
University of Zagreb
Pierottijeva 6, 10000 Zagreb
Croatia
pecaric@hazu.hr

Abstract. Some new extensions of the weighted Montgomery identity are given, and used to generalize the Ostrowski inequality and the estimates of the difference between two integral means for α-L-Hölder type functions.

Date: November 30, 2003
2000 *Mathematics Subject Classification.* 26D15, 26D20
Key words and phrases. Taylor's formula, weighted Montgomery identity

1. Introduction

Let $f : [a,b] \to \mathbb{R}$ be differentiable on $[a,b]$, and $f' : [a,b] \to \mathbb{R}$ integrable on $[a,b]$. Then the Montgomery identity holds [7]

$$f(x) = \frac{1}{b-a}\int_a^b f(t)\,dt + \int_a^b P(x,t)\,f'(t)\,dt$$

where $P(x,t)$ is the Peano kernel, defined by

$$P(x,t) = \begin{cases} \frac{t-a}{b-a}, & a \le t \le x, \\ \frac{t-b}{b-a} & x < t \le b. \end{cases} \tag{1.1}$$

Now, let's suppose $w : [a,b] \to [0,\infty\rangle$ is some probability density function, i.e. integrable function satisfying $\int_a^b w(t)\,dt = 1$, and $W(t) = \int_a^t w(x)\,dx$ for $t \in [a,b]$, $W(t) = 0$ for $t < a$ and $W(t) = 1$ for $t > b$. The following identity, given by Pečarić in [8] (see also [2]) is the weighted generalization of Montgomery identity

$$f(x) = \int_a^b w(t)\,f(t)\,dt + \int_a^b P_w(x,t)\,f'(t)\,dt \tag{1.2}$$

where the weighted Peano kernel is

$$P_w(x,t) = \begin{cases} W(t), & a \le t \le x, \\ W(t) - 1 & x < t \le b, \end{cases} \tag{1.3}$$

In the paper [1] G. A. Anastassiou proved the following equality

$$g(y)-g(x)-\sum_{j=1}^{n}\frac{g^{(j)}(x)}{j!}(y-x)^{j}=\frac{1}{(n-1)!}\int_{x}^{y}\left(g^{(n)}(t)-g^{(n)}(x)\right)(y-t)^{n-1}\,dt \tag{1.4}$$

where $g: I \to \mathbb{R}$ is such that $g^{(n)}$ exists for all $t \in [a,b]$, for some $n \in \mathbb{N}$ and $I \subset \mathbb{R}$ is an open interval, $a, b \in I$, $a < b$ and $x, y \in [a,b]$.

In this paper we will use the formula (1.4) to obtain two extensions of weighted Montgomery identity (Section 2.) and further to obtain some new Ostrowski type inequalities (Section 3.), as well as some generalizations of the estimations of the difference of two weighted integral means (Section 4.).

2. Two extensions of Montgomery identity

Theorem 1. *Let $f: I \to \mathbb{R}$ be such that $f^{(n-1)}$ is absolutely continuous for some $n \geq 2$, $I \subset \mathbb{R}$ an open interval, $a, b \in I$, $a < b$ and $w: [a,b] \to [0,\infty\rangle$ some probability density function. Then the following identity hold*

$$\begin{aligned} f(x) &= \int_{a}^{b} w(t) f(t)\,dt+\sum_{j=0}^{n-1}\frac{f^{(j+1)}(a)}{j!}\int_{a}^{x} W(t)(t-a)^{j}\,dt \\ &+\sum_{j=0}^{n-1}\frac{f^{(j+1)}(b)}{j!}\int_{x}^{b}(W(t)-1)(t-b)^{j}\,dt \\ &+\frac{1}{(n-2)!}\int_{a}^{x} W(t)\left[\int_{a}^{t}\left(f^{(n)}(s)-f^{(n)}(a)\right)(t-s)^{n-2}\,ds\right]dt \\ &+\frac{1}{(n-2)!}\int_{x}^{b}(1-W(t))\left[\int_{t}^{b}\left(f^{(n)}(s)-f^{(n)}(b)\right)(t-s)^{n-2}\,ds\right]dt \end{aligned} \tag{2.1}$$

Proof. If we apply formula (1.4) with f' instead of g, and with $x = a$ or $x = b$, then replace n with $n-1$ (thus $n \geq 2$), we have

$$f'(t)=f'(a)+\sum_{j=1}^{n-1}\frac{f^{(j+1)}(a)}{j!}(t-a)^{j}+\frac{1}{(n-2)!}\int_{a}^{t}\left(f^{(n)}(s)-f^{(n)}(a)\right)(t-s)^{n-2}\,ds,$$

$$f'(t)=f'(b)+\sum_{j=1}^{n-1}\frac{f^{(j+1)}(b)}{j!}(t-b)^{j}-\frac{1}{(n-2)!}\int_{t}^{b}\left(f^{(n)}(s)-f^{(n)}(b)\right)(t-s)^{n-2}\,ds.$$

By putting these two formulae in the weighted Montgomery identity (1.2) we obtain (2.1)

$$\begin{aligned}
f(x) &= \int_a^b w(t) f(t)\, dt + f'(a) \int_a^x W(t)\, dt + \sum_{j=1}^{n-1} \frac{f^{(j+1)}(a)}{j!} \int_a^x W(t)(t-a)^j\, dt \\
&+ f'(b) \int_x^b (W(t)-1)\, dt + \sum_{j=1}^{n-1} \frac{f^{(j+1)}(b)}{j!} \int_x^b (W(t)-1)(t-b)^j\, dt \\
&+ \frac{1}{(n-2)!} \int_a^x W(t) \left[\int_a^t \left(f^{(n)}(s) - f^{(n)}(a) \right) (t-s)^{n-2}\, ds \right] dt \\
&+ \frac{1}{(n-2)!} \int_x^b (1-W(t)) \left[\int_t^b \left(f^{(n)}(s) - f^{(n)}(b) \right) (t-s)^{n-2}\, ds \right] dt.
\end{aligned}$$

□

Theorem 2. *Let $f : I \to \mathbb{R}$ be such that $f^{(n-1)}$ is absolutely continuous for some $n \geq 2$, $I \subset \mathbb{R}$ an open interval, $a, b \in I$, $a < b$ and $w : [a,b] \to [0,\infty\rangle$ some probability density function. Then the following identity hold*

$$\begin{aligned}
f(x) &= \int_a^b w(t) f(t)\, dt + \sum_{j=0}^{n-1} \frac{f^{(j+1)}(x)}{j!} \int_a^b P_w(x,t)(t-x)^j\, dt \\
&+ \frac{1}{(n-2)!} \int_a^b P_w(x,t) \left[\int_x^t \left(f^{(n)}(s) - f^{(n)}(x) \right) (t-s)^{n-2}\, ds \right] dt.
\end{aligned} \tag{2.2}$$

Proof. If we apply formula (1.4) with f' instead of g, then replace n with $n-1$ (thus $n \geq 2$), we have

$$f'(t) = f'(x) + \sum_{j=1}^{n-1} \frac{f^{(j+1)}(x)}{j!} (t-x)^j + \frac{1}{(n-2)!} \int_x^t \left(f^{(n)}(s) - f^{(n)}(x) \right) (t-s)^{n-2}\, ds.$$

By putting this formula in the weighted Montgomery identity (1.2) we obtain (2.2)

$$\begin{aligned}
f(x) &= \int_a^b w(t) f(t)\, dt + f'(x) \int_a^b P_w(x,t)\, dt + \sum_{j=1}^{n-1} \frac{f^{(j+1)}(x)}{j!} \int_a^b P_w(x,t)(t-x)^j\, dt \\
&+ \frac{1}{(n-2)!} \int_a^b P_w(x,t) \left[\int_x^t \left(f^{(n)}(s) - f^{(n)}(x) \right) (t-s)^{n-2}\, ds \right] dt.
\end{aligned}$$

□

Remark 1. *In the special case, if we take $w(t)=\frac{1}{b-a}$, $t\in[a,b]$ the equality (2.1) reduces to*

$$\begin{aligned} f(x) &= \frac{1}{b-a}\int_a^b f(t)\,dt+\sum_{j=0}^{n-1}\frac{f^{(j+1)}(a)}{j!\,(j+2)}(x-a)^{j+2}-\sum_{j=0}^{n-1}\frac{f^{(j+1)}(b)}{j!\,(j+2)}(x-b)^{j+2} \\ &\quad+\frac{1}{(n-2)!}\int_a^x \frac{t-a}{b-a}\left[\int_a^t \left(f^{(n)}(s)-f^{(n)}(a)\right)(t-s)^{n-2}\,ds\right]dt \\ &\quad+\frac{1}{(n-2)!}\int_x^b \frac{b-t}{b-a}\left[\int_t^b \left(f^{(n)}(s)-f^{(n)}(b)\right)(t-s)^{n-2}\,ds\right]dt \end{aligned} \tag{2.3}$$

and the equality (2.2) reduces to

$$\begin{aligned} f(x) &= \frac{1}{b-a}\int_a^b f(t)\,dt+\sum_{j=0}^{n-1}\frac{f^{(j+1)}(x)}{(b-a)\,(j+2)!}\left((a-x)^{j+2}-(b-x)^{j+2}\right) \\ &\quad+\frac{1}{(n-2)!}\int_a^b P(x,t)\left[\int_x^t \left(f^{(n)}(s)-f^{(n)}(x)\right)(t-s)^{n-2}\,ds\right]dt \end{aligned} \tag{2.4}$$

where $P(x,t)$ is given by (1.1).

3. The Ostrowski type inequalities

In this section we generalize the results from [1], [5] and [6]. We denote

$$v_{w,n}^{[a,b]}(x)=\sum_{j=0}^{n-1}\frac{f^{(j+1)}(a)}{j!}\int_a^x W(t)(t-a)^j\,dt+\sum_{j=0}^{n-1}\frac{f^{(j+1)}(b)}{j!}\int_x^b (W(t)-1)(t-b)^j\,dt,$$

$$t_{w,n}^{[a,b]}(x)=\sum_{j=0}^{n-1}\frac{f^{(j+1)}(x)}{j!}\int_a^b P_w(x,t)(t-x)^j\,dt.$$

The Beta and the incomplete Beta function of Euler type are defined by

$$B(x,y)=\int_0^1 t^{x-1}(1-t)^{y-1}\,dt,\qquad B_r(x,y)=\int_0^r t^{x-1}(1-t)^{y-1}\,dt,\qquad x,y>0$$

and

$$\Psi_r(x,y)=\int_0^r t^{x-1}(1+t)^{y-1}\,dt$$

is a real positive valued integral.

Theorem 3. *Suppose that all the assumptions of Theorem 1 hold. Additionally assume that $f^{(n)}:[a,b]\to\mathbb{R}$ is an α-L-Hölder type function, i.e. $\left|f^{(n)}(x)-f^{(n)}(y)\right|\leq$*

$L|x-y|^{\alpha}$ *for every* $x,y\in[a,b]$, *where* $L>0$ *and* $\alpha\in\langle 0,1]$. *Then we have*

$$\begin{aligned}
&\left|f(x)-\int_a^b w(t)f(t)\,dt-v_{w,n}^{[a,b]}(x)\right| \\
&\le \frac{B(\alpha+1,n-1)}{(n-2)!}L\left(\int_a^x |W(t)|(t-a)^{\alpha+n-1}\,dt+\int_x^b |W(t)-1|(b-t)^{\alpha+n-1}\,dt\right) \\
&\le \frac{B(\alpha+1,n-1)}{(n-2)!(\alpha+n)}L\left[(b-x)^{\alpha+n}+(x-a)^{\alpha+n}\right] \qquad (3.1)
\end{aligned}$$

Proof. We use the identity (2.1) and apply the properties of modulus to obtain

$$\begin{aligned}
&\left|f(x)-\int_a^b w(t)f(t)\,dt-v_{w,n}^{[a,b]}(x)\right| \\
&\le \frac{1}{(n-2)!}\left\{\left|\int_a^x W(t)\left[\int_a^t \left(f^{(n)}(s)-f^{(n)}(a)\right)(t-s)^{n-2}\,ds\right]dt\right|\right. \\
&\left.+\left|\int_x^b (1-W(t))\left[\int_t^b \left(f^{(n)}(s)-f^{(n)}(b)\right)(t-s)^{n-2}\,ds\right]dt\right|\right\} \\
&\le \frac{L}{(n-2)!}\left\{\int_a^x |W(t)|\left|\int_a^t |s-a|^{\alpha}|t-s|^{n-2}\,ds\right|dt\right. \\
&\left.+\int_x^b |1-W(t)|\left|\int_t^b |s-b|^{\alpha}|t-s|^{n-2}\,ds\right|dt\right\}.
\end{aligned}$$

With $(s-a)=u(t-a)$

$$\begin{aligned}
\int_a^t |s-a|^{\alpha}|(t-s)|^{n-2}\,ds &= \int_a^t (s-a)^{\alpha}(t-s)^{n-2}\,ds \\
&= (t-a)^{\alpha+n-1}B(\alpha+1,n-1),
\end{aligned}$$

and with $(b-s)=u(b-t)$

$$\begin{aligned}
\int_t^b |s-b|^{\alpha}|t-s|^{n-2}\,ds &= \int_t^b (b-s)^{\alpha}(s-t)^{n-2}\,ds \\
&= (b-t)^{\alpha+n-1}B(\alpha+1,n-1)
\end{aligned}$$

and the first inequality from (3.2) follows. Since $0\le W(t)\le 1$, $t\in[a,b]$ and $0\le 1-W(t)\le 1$, $t\in[a,b]$, so we have

$$\int_a^x |W(t)|(t-a)^{\alpha+n-1}\,dt \le \int_a^x (t-a)^{\alpha+n-1}\,dt=\frac{(x-a)^{\alpha+n}}{\alpha+n},$$

$$\int_x^b |W(t)-1|(b-t)^{\alpha+n-1}\,dt \le \int_x^b (b-t)^{\alpha+n-1}\,dt=\frac{(b-x)^{\alpha+n}}{\alpha+n}.$$

The second inequality from (3.2) follows. □

Theorem 4. *Suppose that all the assumptions of Theorem 2 hold. Additionally assume that* $f^{(n)}:[a,b]\to\mathbb{R}$ *is an* α-L-*Hölder type function, i.e.* $\left|f^{(n)}(x)-f^{(n)}(y)\right|\le$

$L\left|x-y\right|^{\alpha}$ *for every* $x,y\in[a,b]$, *where* $L>0$ *and* $\alpha\in\langle 0,1]$. *Then we have*

$$\begin{aligned}&\left|f\left(x\right)-\int_{a}^{b}w\left(t\right)f\left(t\right)dt-t_{w,n}^{[a,b]}\left(x\right)\right|\\&\leq\frac{B\left(\alpha+1,n-1\right)}{(n-2)!}L\int_{a}^{b}\left|P_{w}\left(x,t\right)\right|\left|t-x\right|^{\alpha+n-1}dt\\&\leq\frac{B\left(\alpha+1,n-1\right)}{(n-2)!\left(\alpha+n\right)}L\left[\left(b-x\right)^{\alpha+n}+\left(x-a\right)^{\alpha+n}\right]\end{aligned}\tag{3.2}$$

Proof. We use the identity (2.2) and apply the properties of modulus to obtain

$$\begin{aligned}&\left|f\left(x\right)-\int_{a}^{b}w\left(t\right)f\left(t\right)dt-t_{w,n}^{[a,b]}\left(x\right)\right|\\&=\left|\frac{1}{(n-2)!}\int_{a}^{b}P_{w}\left(x,t\right)\left[\int_{x}^{t}\left(f^{(n)}\left(s\right)-f^{(n)}\left(x\right)\right)\left(t-s\right)^{n-2}ds\right]dt\right|\\&\leq\frac{1}{(n-2)!}\int_{a}^{b}\left|P_{w}\left(x,t\right)\right|\left|\int_{x}^{t}\left|\left(f^{(n)}\left(s\right)-f^{(n)}\left(x\right)\right)\left(t-s\right)^{n-2}\right|ds\right|dt\\&\leq\frac{L}{(n-2)!}\int_{a}^{b}\left|P_{w}\left(x,t\right)\right|\left|\int_{x}^{t}\left|s-x\right|^{\alpha}\left|\left(t-s\right)\right|^{n-2}ds\right|dt.\end{aligned}$$

Now, for $t>x$ let $(s-x)=u\left(t-x\right)$

$$\begin{aligned}\int_{x}^{t}\left|s-x\right|^{\alpha}\left|\left(t-s\right)\right|^{n-2}ds&=\int_{x}^{t}\left(s-x\right)^{\alpha}\left(t-s\right)^{n-2}ds\\&=\left(t-x\right)^{\alpha+n-1}B\left(\alpha+1,n-1\right),\end{aligned}$$

for $x<t$ let $(x-s)=u\left(x-t\right)$

$$\begin{aligned}\int_{x}^{t}\left|s-x\right|^{\alpha}\left|\left(t-s\right)\right|^{n-2}ds&=-\int_{t}^{x}\left(s-x\right)^{\alpha}\left(t-s\right)^{n-2}ds\\&=-\left(x-t\right)^{\alpha+n-1}B\left(\alpha+1,n-1\right)\end{aligned}$$

so

$$\left|\int_{x}^{t}\left|s-x\right|^{\alpha}\left|\left(t-s\right)\right|^{n-2}ds\right|=\left|t-x\right|^{\alpha+n-1}B\left(\alpha+1,n-1\right)\tag{3.3}$$

and the first inequality from (3.2) follows. Since $0\leq W\left(t\right)\leq 1$, $t\in[a,b]$ then $\left|P_{w}\left(x,t\right)\right|\leq 1$, $t\in[a,b]$, so we have

$$\int_{a}^{b}\left|P_{w}\left(x,t\right)\right|\left|t-x\right|^{\alpha+n-1}dt\leq\int_{a}^{b}\left|t-x\right|^{\alpha+n-1}dt=\frac{\left(b-x\right)^{\alpha+n}+\left(x-a\right)^{\alpha+n}}{\alpha+n}.$$

The second inequality from (3.2) follows. □

Corollary 1. *Suppose that all the assumptions of Theorem 3 hold. Then we have*

$$\begin{aligned}&\left|f\left(x\right)-\frac{1}{b-a}\int_{a}^{b}f\left(t\right)dt-\sum_{j=0}^{n-1}\frac{f^{(j+1)}\left(a\right)}{j!\left(j+2\right)}\left(x-a\right)^{j+2}+\sum_{j=0}^{n-1}\frac{f^{(j+1)}\left(b\right)}{j!\left(j+2\right)}\left(x-b\right)^{j+2}\right|\\&\leq B\left(\alpha+1,n-1\right)L\frac{\left(x-a\right)^{\alpha+n+1}+\left(b-x\right)^{\alpha+n+1}}{\left(\alpha+n+1\right)\left(b-a\right)\left(n-2\right)!}.\end{aligned}$$

Proof. We apply the first inequality from (3.1) with $w(t) = \frac{1}{b-a}$, $t \in [a,b]$

$$\begin{aligned}
&\int_a^x |W(t)| (t-a)^{\alpha+n-1} dt + \int_x^b |W(t) - 1| (b-t)^{\alpha+n-1} dt \\
&= \frac{1}{b-a} \left(\int_a^x |t-a| (t-a)^{\alpha+n-1} dt + \int_x^b |t-b| (b-t)^{\alpha+n-1} dt \right) \\
&= \frac{(x-a)^{\alpha+n+1} + (b-x)^{\alpha+n+1}}{(b-a)(\alpha+n+1)}
\end{aligned}$$

and the inequality follows from the Theorem 3. □

Corollary 2. *Suppose that all the assumptions of Theorem 4 hold. Then we have*

$$\begin{aligned}
&\left| f(x) - \frac{1}{b-a} \int_a^b f(t)\, dt - \sum_{j=0}^{n-1} \frac{f^{(j+1)}(x)}{(b-a)(j+2)!} \left((a-x)^{j+2} - (b-x)^{j+2} \right) \right| \\
&\leq \frac{B(\alpha+1, n-1)\, B(2, \alpha+n)}{(b-a)(n-2)!} L \left[(x-a)^{\alpha+n+1} + (b-x)^{\alpha+n+1} \right].
\end{aligned}$$

Proof. We apply the first inequality from (3.2) with $w(t) = \frac{1}{b-a}$, $t \in [a,b]$

$$\begin{aligned}
&\int_a^b |P(x,t)| \, |t-x|^{\alpha+n-1} dt \\
&= \frac{1}{b-a} \left(\int_a^x |(t-a)| \, |t-x|^{\alpha+n-1} dt + \int_x^b |(t-b)| \, |t-x|^{\alpha+n-1} dt \right).
\end{aligned}$$

Using substitution $t - a = u(x-a)$ the first integral is equal to

$$\begin{aligned}
\int_a^x (t-a)(x-t)^{\alpha+n-1} dt &= (x-a)^{\alpha+n+1} \int_0^1 u(1-u)^{\alpha+n-1} du \\
&= (x-a)^{\alpha+n+1} B(2, \alpha+n)
\end{aligned}$$

and similarly using $b - t = u(b-x)$ the second is

$$\int_x^b (b-t)(t-x)^{\alpha+n-1} dt = (b-x)^{\alpha+n+1} B(2, \alpha+n).$$

Finally

$$\int_a^b |P(x,t)| \, |t-x|^{\alpha+n-1} dt = \frac{(x-a)^{\alpha+n+1} + (b-x)^{\alpha+n+1}}{b-a} B(2, \alpha+n).$$

and the inequality follows from the Theorem 4. □

Remark 2. *If we apply (3.1) and (3.2) with* $x = \frac{a+b}{2}$ *we get the generalized mid-point inequalities*

$$\begin{aligned}&\left|f\left(\frac{a+b}{2}\right)-\int_a^b w(t)\,f(t)\,dt-v_{w,n}^{[a,b]}\left(\frac{a+b}{2}\right)\right| \\ &\leq \frac{B(\alpha+1,n-1)}{(n-2)!}L\left(\int_a^{\frac{a+b}{2}}|W(t)|\,(t-a)^{\alpha+n-1}\,dt+\int_{\frac{a+b}{2}}^b|W(t)-1|\,(b-t)^{\alpha+n-1}\,dt\right) \\ &\leq \frac{B(\alpha+1,n-1)}{(n-2)!\,(\alpha+n)}L\left[\frac{(b-a)^{\alpha+n}}{2^{\alpha+n-1}}\right]\end{aligned}$$

and

$$\begin{aligned}&\left|f\left(\frac{a+b}{2}\right)-\int_a^b w(t)\,f(t)\,dt-t_{w,n}^{[a,b]}\left(\frac{a+b}{2}\right)\right| \\ &\leq \frac{B(\alpha+1,n-1)}{(n-2)!}L\int_a^b\left|P_w\left(\frac{a+b}{2},t\right)\right|\left|t-\frac{a+b}{2}\right|^{\alpha+n-1}dt \\ &\leq \frac{B(\alpha+1,n-1)}{(n-2)!\,(\alpha+n)}L\left[\frac{(b-a)^{\alpha+n}}{2^{\alpha+n-1}}\right].\end{aligned}$$

If we additionally assume that $w(t)$ *is symmetric on* $[a,b]$ *i.e.* $w(t)=w(b-a-t)$ *for every* $t\in[a,b]$ *these inequalities reduce to*

$$\begin{aligned}&\left|f\left(\frac{a+b}{2}\right)-\int_a^b w(t)\,f(t)\,dt\right. \\ &\left.-\sum_{j=0}^{n-1}\frac{f^{(j+1)}(a)+(-1)^{j+1}f^{(j+1)}(b)}{j!}\int_a^{\frac{a+b}{2}}W(t)\,(t-a)^j\,dt\right| \\ &\leq \frac{B(\alpha+1,n-1)}{(n-2)!}2L\int_a^{\frac{a+b}{2}}W(t)\,(t-a)^{\alpha+n-1}\,dt \\ &\leq \frac{B(\alpha+1,n-1)}{(n-2)!\,(\alpha+n)}L\left[\frac{(b-a)^{\alpha+n}}{2^{\alpha+n-1}}\right]\end{aligned}$$

and

$$\begin{aligned}&\left|f\left(\frac{a+b}{2}\right)-\int_a^b w(t)\,f(t)\,dt-\sum_{j=0}^{\lfloor\frac{n-1}{2}\rfloor}\frac{f^{(2j+1)}(x)}{(2j)!}2\int_a^{\frac{a+b}{2}}W(t)\left(t-\frac{a+b}{2}\right)^j dt\right| \\ &\leq \frac{B(\alpha+1,n-1)}{(n-2)!}2L\int_a^{\frac{a+b}{2}}W(t)\left(\frac{a+b}{2}-t\right)^{\alpha+n-1}dt \\ &\leq \frac{B(\alpha+1,n-1)}{(n-2)!\,(\alpha+n)}L\left[\frac{(b-a)^{\alpha+n}}{2^{\alpha+n-1}}\right].\end{aligned}$$

For the generalized trapezoid inequality we apply equality (2.1) and (2.2) first with $x=a$*, then with* $x=b$*, then add them up and divide by* 2. *After applying the*

properties of modulus we obtain

$$\begin{aligned}
&\left|\frac{f(a)+f(b)}{2}-\int_a^b w(t)\,f(t)\,dt-\frac{v_{w,n}^{[a,b]}(a)+v_{w,n}^{[a,b]}(b)}{2}\right| \\
&\leq \frac{B(\alpha+1,n-1)}{(n-2)!}L\left(\int_a^b W(t)\,(t-a)^{\alpha+n-1}\,dt+\int_a^b (1-W(t))\,(b-t)^{\alpha+n-1}\,dt\right) \\
&\leq \frac{B(\alpha+1,n-1)}{(n-2)!\,(\alpha+n)}2L\left[(b-a)^{\alpha+n}\right]
\end{aligned}$$

and

$$\begin{aligned}
&\left|\frac{f(a)+f(b)}{2}-\int_a^b w(t)\,f(t)\,dt-\frac{t_{w,n}^{[a,b]}(a)+t_{w,n}^{[a,b]}(b)}{2}\right| \\
&\leq \frac{B(\alpha+1,n-1)}{(n-2)!}L\left(\int_a^b (1-W(t))\,(t-a)^{\alpha+n-1}\,dt+\int_a^b W(t)\,(b-t)^{\alpha+n-1}\,dt\right) \\
&\leq \frac{B(\alpha+1,n-1)}{(n-2)!\,(\alpha+n)}2L\left[(b-a)^{\alpha+n}\right].
\end{aligned}$$

If we additionally assume that $w(t)$ *is symmetric on* $[a,b]$*, these inequalities reduce to*

$$\begin{aligned}
&\left|\frac{f(a)+f(b)}{2}-\int_a^b w(t)\,f(t)\,dt\right. \\
&\left.-\sum_{j=0}^{n-1}\frac{f^{(j+1)}(a)+(-1)^{j+1}f^{(j+1)}(b)}{2\,(j!)}\int_a^b W(t)\,(b-a)^j\,dt\right| \\
&\leq \frac{B(\alpha+1,n-1)}{(n-2)!}2L\left(\int_a^b W(t)\,(t-a)^{\alpha+n-1}\,dt\right) \\
&\leq \frac{B(\alpha+1,n-1)}{(n-2)!\,(\alpha+n)}2L\left[(b-a)^{\alpha+n}\right]
\end{aligned}$$

and

$$\begin{aligned}
&\left|\frac{f(a)+f(b)}{2}-\int_a^b w(t)\,f(t)\,dt\right. \\
&\left.-\sum_{j=0}^{n-1}\frac{f^{(j+1)}(a)}{2\,(j!)}\int_a^b (W(t)-1)\,(t-a)^j\,dt-\sum_{j=0}^{n-1}\frac{f^{(j+1)}(b)}{2\,(j!)}\int_a^b W(t)\,(t-b)^j\,dt\right| \\
&\leq \frac{B(\alpha+1,n-1)}{(n-2)!}2L\left(\int_a^b W(t)\,(t-a)^{\alpha+n-1}\,dt\right) \\
&\leq \frac{B(\alpha+1,n-1)}{(n-2)!\,(\alpha+n)}2L\left[(b-a)^{\alpha+n}\right].
\end{aligned}$$

4. The estimates of the difference between the two weighted integral means

In this section we generalize the results from [3], [4]. For the two intervals $[a,b]$ and $[c,d]$ we have four possible cases if $[a,b]\cap[c,d]\neq\emptyset$. The first case is $[c,d]\subset[a,b]$ and the second $[a,b]\cap[c,d]=[c,b]$. Other two possible cases we simply get by change $a\leftrightarrow c$, $b\leftrightarrow d$.

Theorem 5. *Let $f:I\to\mathbb{R}$ be such that $[a,b]\cup[c,d]\subset I$, $f^{(n)}:[a,b]\cup[c,d]\to\mathbb{R}$ is an α-L-Hölder type function for some $n\geq 2$, $w:[a,b]\to[0,\infty\rangle$ and $u:[c,d]\to[0,\infty\rangle$ some probability density functions, $W(t)=\int_a^t w(x)\,dx$ for $t\in[a,b]$, $W(t)=0$ for $t<a$ and $W(t)=1$ for $t>b$, $U(t)=\int_c^t u(x)\,dx$ for $t\in[c,d]$, $U(t)=0$ for $t<c$ and $U(t)=1$ for $t>d$. Then if $[a,b]\cap[c,d]\neq\emptyset$ and $x\in[a,b]\cap[c,d]$, we have*

$$\int_a^b w(t)f(t)\,dt-\int_c^d u(t)f(t)\,dt+t_{w.n}^{[a,b]}(x)-t_{u,n}^{[c,d]}(x)=\int_{\min\{a,c\}}^{\max\{b,d\}}K_n(x,t)\,dt \tag{4.1}$$

where in case $[c,d]\subset[a,b]$

$$K_n(x,t)=\begin{cases}\frac{-1}{(n-2)!}W(t)\left[\int_x^t\left(f^{(n)}(s)-f^{(n)}(x)\right)(t-s)^{n-2}\,ds\right], & t\in[a,c\rangle,\\ \frac{-1}{(n-2)!}\left(W(t)-U(t)\right)\left[\int_x^t\left(f^{(n)}(s)-f^{(n)}(x)\right)(t-s)^{n-2}\,ds\right], & t\in[c,d],\\ \frac{-1}{(n-2)!}\left(W(t)-1\right)\left[\int_x^t\left(f^{(n)}(s)-f^{(n)}(x)\right)(t-s)^{n-2}\,ds\right], & t\in\langle d,b],\end{cases}$$

and in case $[a,b]\cap[c,d]=[c,b]$

$$K_n(x,t)=\begin{cases}\frac{-1}{(n-2)!}W(t)\left[\int_x^t\left(f^{(n)}(s)-f^{(n)}(x)\right)(t-s)^{n-2}\,ds\right], & t\in[a,c\rangle,\\ \frac{-1}{(n-2)!}\left(W(t)-U(t)\right)\left[\int_x^t\left(f^{(n)}(s)-f^{(n)}(x)\right)(t-s)^{n-2}\,ds\right], & t\in[c,b],\\ \frac{1}{(n-2)!}\left(U(t)-1\right)\left[\int_x^t\left(f^{(n)}(s)-f^{(n)}(x)\right)(t-s)^{n-2}\,ds\right], & t\in\langle b,d].\end{cases}$$

Proof. We subtract identities (2.2) for interval $[a,b]$ and $[c,d]$, to get the formula (4.1). □

Theorem 6. *Suppose that all the assumptions of Theorem 5 hold. Then we have*

$$\begin{aligned}&\left|\int_a^b w(t)f(t)\,dt-\int_c^d u(t)f(t)\,dt+t_{w,n}^{[a,b]}(x)-t_{u,n}^{[c,d]}(x)\right|\\ &\leq\frac{B(\alpha+1,n-1)}{(n-2)!}L\int_{\min\{a,c\}}^{\max\{b,d\}}\left|P_w(x,t)-P_u(x,t)\right||t-x|^{\alpha+n-1}\,dt\\ &\leq\frac{B(\alpha+1,n-1)}{(n-2)!}L\int_{\min\{a,c\}}^{\max\{b,d\}}|t-x|^{\alpha+n-1}\,dt \end{aligned}\tag{4.2}$$

Proof. Use the identity (4.1) to obtain

$$\left|\int_a^b w(t) f(t)\,dt - \int_c^d u(t) f(t)\,dt + t_{w,n}^{[a,b]}(x) - t_{u,n}^{[c,d]}(x)\right| \leq \int_{\min\{a,c\}}^{\max\{b,d\}} |K_n(x,t)|\,dt$$
$$\leq \frac{1}{(n-2)!}\left(\int_{\min\{a,c\}}^{\max\{b,d\}} |P_w(x,t) - P_u(x,t)| \left|\int_x^t \left|f^{(n)}(s) - f^{(n)}(x)\right| |t-s|^{n-2}\,ds\right| dt\right)$$
$$\leq \frac{L}{(n-2)!}\left(\int_{\min\{a,c\}}^{\max\{b,d\}} |P_w(x,t) - P_u(x,t)| \left|\int_x^t |s-x|^{\alpha} |t-s|^{n-2}\,ds\right| dt\right)$$

and like (3.3) in the proof of Theorem 4

$$\left|\int_x^t |s-x|^{\alpha} |t-s|^{n-2}\,ds\right| = |t-x|^{\alpha+n-1} B(\alpha+1, n-1)$$

which proves the first inequality. The second follows from properties $|P_w(x,t)| \leq 1$, $|P_u(x,t)| \leq 1$, $|P_w(x,t) - P_u(x,t)| \leq 1$ for all $t \in [\min\{a,c\}, \max\{b,d\}]$. □

4.1. **Case** $[c,d] \subset [a,b]$**.** Here we denote

$$t_n^{[a,b]}(x) = \sum_{j=0}^{n-1} \frac{f^{(j+1)}(x)}{(b-a)(j+2)!}\left((a-x)^{j+2} - (b-x)^{j+2}\right).$$

Corollary 3. *Suppose that all the assumptions of Theorem 5 hold. Additionally suppose* $[c,d] \subset [a,b]$*. For* $x \in [c,d]$ *and* $s_0 = \frac{bc-ad}{c-a+b-d}$*, the following inequality holds*

$$\left|\frac{1}{b-a}\int_a^b f(t)\,dt - \frac{1}{d-c}\int_c^d f(t)\,dt + t_n^{[a,b]}(x) - t_n^{[c,d]}(x)\right|$$
$$\leq \frac{B(\alpha+1,n-1)}{(b-a)(n-2)!} L\left[(x-a)^{\alpha+n+1} B_{\frac{c-a}{x-a}}(2,\alpha+n) + \frac{(c-a+b-d)}{(d-c)}|x-s_o|^{\alpha+n-1}\right.$$
$$\left.\cdot\left(B(2,\alpha+n) + \Psi_{r_1}(2,\alpha+n) + \Psi_{r_2}(\alpha+n,2)\right) + (b-x)^{\alpha+n+1} B_{\frac{b-d}{b-x}}(2,\alpha+n)\right]$$

where for $s_0 < x$

$$r_1 = \frac{s_0 - c}{x - s_0}, \qquad r_2 = \frac{d-x}{x-s_0},$$

while for $s_0 > x$

$$r_1 = \frac{d - s_0}{s_0 - x}, \qquad r_2 = \frac{x-c}{s_0-x}.$$

Proof. We put $w(t) = \frac{1}{b-a}$, $t \in [a,b]$, and $u(t) = \frac{1}{d-c}$, $t \in [c,d]$ in the first inequality from the Theorem 6. Thus we have $t_n^{[a,b]}(x)$ and $t_n^{[c,d]}(x)$ instead of $t_{w,n}^{[a,b]}(x)$ and $t_{u,n}^{[c,d]}(x)$ and

$$\int_{\min\{a,c\}}^{\max\{b,d\}} |P_w(x,t) - P_u(x,t)| |t-x|^{\alpha+n-1}\,dt$$
$$= \int_a^c \left|\frac{t-a}{b-a}\right| |x-t|^{\alpha+n-1}\,dt + \int_c^d \left|\frac{t-a}{b-a} - \frac{t-c}{d-c}\right| |x-t|^{\alpha+n-1}\,dt$$
$$+ \int_d^b \left|\frac{b-t}{b-a}\right| |x-t|^{\alpha+n-1}\,dt.$$

For the first integral let $t-a=u(x-a)$ so

$$\int_a^c (t-a)(x-t)^{\alpha+n-1}\,dt = (x-a)^{\alpha+n+1}\int_0^{\frac{c-a}{x-a}} u(1-u)^{\alpha+n-1}\,du$$
$$= (x-a)^{\alpha+n+1} B_{\frac{c-a}{x-a}}(2,\alpha+n). \tag{4.3}$$

For the third integral let $b-t=u(b-x)$ and

$$\int_d^b (b-t)(t-x)^{\alpha+n-1}\,dt = -(b-x)^{\alpha+n+1}\int_{\frac{b-d}{b-x}}^{0} u(1-u)^{\alpha+n-1}\,du$$
$$= (b-x)^{\alpha+n+1} B_{\frac{b-d}{b-x}}(2,\alpha+n).$$

$c-a+b-d>0$ so the second integral is

$$\int_c^d \left|\frac{t-a}{b-a}-\frac{t-c}{d-c}\right| |x-t|^{\alpha+n-1}\,dt$$
$$= \frac{c-a+b-d}{(b-a)(d-c)}\int_c^d |s_0-t|\,|x-t|^{\alpha+n-1}\,dt.$$

Since $s_0-c=\frac{(d-c)(c-a)}{c-a+b-d}\geq 0$ and $d-s_0=\frac{(d-c)(b-d)}{c-a+b-d}\geq 0$, $s_0\in[c,d]$. So we have two possible cases:

1. If $s_0<x$ we have

$$\int_c^d |s_0-t|\,|x-t|^{\alpha+n-1}\,dt$$
$$= \int_c^{s_0}(s_0-t)(x-t)^{\alpha+n-1}\,dt + \int_{s_0}^{x}(t-s_0)(x-t)^{\alpha+n-1}\,dt$$
$$+ \int_x^d (t-s_0)(t-x)^{\alpha+n-1}\,dt.$$

Now, using the substitution $s_0-t=u(x-s_0)$ we get

$$\int_c^{s_0}(s_0-t)(x-t)^{\alpha+n-1}\,dt = -(x-s_0)^{\alpha+n-1}\int_{\frac{s_0-c}{x-s_0}}^{0} u(1+u)^{\alpha+n-1}\,du$$
$$= (x-s_0)^{\alpha+n-1}\,\Psi_{\frac{s_0-c}{x-s_0}}(2,\alpha+n),$$

with $t-s_0=u(x-s_0)$ we get

$$\int_{s_0}^{x}(t-s_0)(x-t)^{\alpha+n-1}\,dt = (x-s_0)^{\alpha+n-1}\int_0^1 u(1-u)^{\alpha+n-1}\,du$$
$$= (x-s_0)^{\alpha+n-1} B(2,\alpha+n),$$

and with $t-x=u(x-s_0)$

$$\int_x^d (t-s_0)(t-x)^{\alpha+n-1}\,dt = (x-s_0)^{\alpha+n-1}\int_0^{\frac{d-x}{x-s_0}} u^{\alpha+n-1}(1+u)\,du$$
$$= (x-s_0)^{\alpha+n-1}\,\Psi_{\frac{d-x}{x-s_0}}(\alpha+n,2).$$

2. If $x < s_0$ then

$$\begin{aligned}
&\int_c^d |s_0 - t| \, |x - t|^{\alpha+n-1} \, dt \\
&= \int_c^x (s_0 - t) (x - t)^{\alpha+n-1} \, dt + \int_x^{s_0} (s_0 - t) (t - x)^{\alpha+n-1} \, dt \\
&+ \int_{s_0}^d (t - s_0) (t - x)^{\alpha+n-1} \, dt.
\end{aligned}$$

Using the substitution $x - t = u (s_0 - x)$ we get

$$\begin{aligned}
\int_c^x (s_0 - t) (x - t)^{\alpha+n-1} \, dt &= - (s_0 - x)^{\alpha+n-1} \int_{\frac{x-c}{s_0-x}}^0 (1 + u) \, u^{\alpha+n-1} du \\
&= (s_0 - x)^{\alpha+n-1} \, \Psi_{\frac{x-c}{s_0-x}} (\alpha + n, 2) , \qquad (4.4)
\end{aligned}$$

with $s_0 - t = u (s_0 - x)$ we get

$$\begin{aligned}
\int_x^{s_0} (s_0 - t) (t - x)^{\alpha+n-1} \, dt &= (s_0 - x)^{\alpha+n-1} \int_0^1 u (1 - u)^{\alpha+n-1} \, du \\
&= (s_0 - x)^{\alpha+n-1} \, B (2, \alpha + n) ,
\end{aligned}$$

and with $t - s_0 = u (s_0 - x)$

$$\begin{aligned}
\int_{s_0}^d (t - s_0) (t - x)^{\alpha+n-1} \, dt &= (s_0 - x)^{\alpha+n-1} \int_0^{\frac{d-s_0}{s_0-x}} u (1 + u)^{\alpha+n-1} \, du \\
&= (s_0 - x)^{\alpha+n-1} \, \Psi_{\frac{d-s_0}{s_0-x}} (2, \alpha + n) .
\end{aligned}$$

Thus the proof is done. □

Remark 3. *If we put $c = d = x$ as a limit case, the inequalities from the Corollary 3 reduce to the inequality from the Corollary 2.*

4.2. **Case** $[a, b] \cap [c, d] = [c, b]$**.**

Corollary 4. *Suppose that all the assumptions of Theorem 5 hold. Additionally suppose $[a, b] \cap [c, d] \subset [c, b]$, $x \in [c, b]$. If $c - a + b - d = 0$, then*

$$\begin{aligned}
&\left| \frac{1}{b-a} \int_a^b f(t) \, dt - \frac{1}{d-c} \int_c^d f(t) \, dt + t_n^{[a,b]} (x) - t_n^{[c,d]} (x) \right| \\
&\leq \frac{B (\alpha + 1, n - 1)}{(n-2)!} L \left[\frac{(x-a)^{\alpha+n+1}}{b-a} B_{\frac{c-a}{x-a}} (2, \alpha + n) \right. \\
&\left. + \frac{c-a}{b-a} \left(\frac{(x-c)^{\alpha+n} + (b-x)^{\alpha+n}}{\alpha + n} \right) + \frac{(d-x)^{\alpha+n+1}}{d-c} B_{\frac{d-b}{d-x}} (2, \alpha + n) \right]
\end{aligned}$$

and if $c-a+b-d\neq 0$ *and* $s_0=\frac{bc-ad}{c-a+b-d}$ *then*

$$\left|\frac{1}{b-a}\int_a^b f(t)\,dt-\frac{1}{d-c}\int_c^d f(t)\,dt+t_n^{[a,b]}(x)-t_n^{[c,d]}(x)\right|$$

$$\leq\frac{B(\alpha+1,n-1)}{(n-2)!}L\left[\frac{(x-a)^{\alpha+n+1}}{b-a}B_{\frac{c-a}{x-a}}(2,\alpha+n)+\frac{|c-a+b-d|}{(b-a)(d-c)}|x-s_o|^{\alpha+n-1}\right.$$

$$\left.\cdot\left(B_{r_1}(\alpha+n,2)+\Psi_{r_2}(\alpha+n,2)\right)+\frac{(d-x)^{\alpha+n+1}}{d-c}B_{\frac{d-b}{d-x}}(2,\alpha+n)\right]$$

where for $s_0<b,c$ *i.e.* $s_0<x$

$$r_1=\frac{x-c}{x-s_0},\qquad r_2=\frac{b-x}{x-s_0},$$

while for $s_0>b,c$ *i.e.* $s_0>x$

$$r_1=\frac{b-x}{s_0-x},\qquad r_2=\frac{x-c}{s_0-x}.$$

Proof. We put $w(t)=\frac{1}{b-a}$, $t\in[a,b]$, and $u(t)=\frac{1}{d-c}$, $t\in[c,d]$ in the first inequality from the Theorem 6. Thus

$$\int_{\min\{a,c\}}^{\max\{b,d\}}|P_w(x,t)-P_u(x,t)|\,|t-x|^{\alpha+n-1}\,dt$$
$$=\int_a^c\left|\frac{t-a}{b-a}\right||x-t|^{\alpha+n-1}\,dt+\int_c^b\left|\frac{t-a}{b-a}-\frac{t-c}{d-c}\right||x-t|^{\alpha+n-1}\,dt$$
$$+\int_b^d\left|\frac{d-t}{d-c}\right||x-t|^{\alpha+n-1}\,dt.$$

The first integral we had before (4.3)

$$\int_a^c(t-a)(x-t)^{\alpha+n-1}\,dt=(x-a)^{\alpha+n+1}B_{\frac{c-a}{x-a}}(2,\alpha+n).$$

For the third integral let $d-t=u(d-x)$ and

$$\int_b^d(d-t)(t-x)^{\alpha+n-1}\,dt=-(d-x)^{\alpha+n+1}\int_{\frac{d-b}{d-x}}^0u(1-u)^{\alpha+n-1}\,du$$
$$=(d-x)^{\alpha+n+1}B_{\frac{d-b}{d-x}}(2,\alpha+n).$$

If $c-a+b-d=0$ then

$$\int_c^b\left|\frac{t-a}{b-a}-\frac{t-c}{d-c}\right||x-t|^{\alpha+n-1}\,dt$$
$$=\frac{c-a}{b-a}\int_c^b|x-t|^{\alpha+n-1}\,dt=\frac{c-a}{b-a}\left(\frac{(x-c)^{\alpha+n}+(b-x)^{\alpha+n}}{\alpha+n}\right)$$

If $c-a+b-d\neq 0$, we have $s_0-c=\frac{(d-c)(c-a)}{c-a+b-d}$ and $s_0-b=\frac{(d-b)(b-a)}{c-a+b-d}$, so there are two possible cases:

1. If $c-a+b-d>0$ (which implies $s_0>b,c$), the second integral is

$$\begin{aligned}
&\int_c^b \left|\frac{t-a}{b-a}-\frac{t-c}{d-c}\right| |x-t|^{\alpha+n-1}\,dt \\
&=\frac{c-a+b-d}{(b-a)(d-c)}\int_c^d |s_0-t|\,|x-t|^{\alpha+n-1}\,dt
\end{aligned}$$

and

$$\int_c^b |s_0-t|\,|x-t|^{\alpha+n-1}\,dt=\int_c^x (s_0-t)(x-t)^{\alpha+n-1}\,dt+\int_x^b (s_0-t)(t-x)^{\alpha+n-1}\,dt.$$

The first integral we had before (4.4)

$$\int_c^x (s_0-t)(x-t)^{\alpha+n-1}\,dt=(s_0-x)^{\alpha+n-1}\,\Psi_{\frac{x-c}{s_0-x}}(\alpha+n,2),$$

and with $t-x=u(s_0-x)$

$$\begin{aligned}
\int_x^b (s_0-t)(t-x)^{\alpha+n-1}\,dt&=(s_0-x)^{\alpha+n-1}\int_0^{\frac{b-x}{s_0-x}}(1-u)\,u^{\alpha+n-1}du \\
&=(s_0-x)^{\alpha+n-1}\,B_{\frac{b-x}{s_0-x}}(\alpha+n,2).
\end{aligned}$$

2. If $c-a+b-d<0$ (which implies $s_0<b,c$) the second integral is

$$\frac{a-c+d-b}{(b-a)(d-c)}\int_c^b |s_0-t|\,|x-t|^{\alpha+n-1}\,dt$$

and

$$\int_c^b |s_0-t|\,|x-t|^{\alpha+n-1}\,dt=\int_c^x (t-s_0)(x-t)^{\alpha+n-1}\,dt+\int_x^b (t-s_0)(t-x)^{\alpha+n-1}\,dt.$$

With $x-t=u(x-s_0)$

$$\begin{aligned}
\int_c^x (t-s_0)(x-t)^{\alpha+n-1}\,dt&=-(s_0-x)^{\alpha+n-1}\int_{\frac{x-c}{x-s_0}}^0 (1-u)\,u^{\alpha+n-1} \\
&=(x-s_0)^{\alpha+n-1}\,B_{\frac{x-c}{x-s_0}}(\alpha+n,2)
\end{aligned}$$

and with $t-x=u(x-s_0)$

$$\int_x^b (s_0-t)(t-x)^{\alpha+n-1}\,dt=(x-s_0)^{\alpha+n-1}\,\Psi_{\frac{b-x}{x-s_0}}(\alpha+n,2).$$

Thus, the proof is done. □

Remark 4. *If we put $b=c=x$ as a limit case, the inequalities from the Corollary 4 reduce to*

$$\begin{aligned}
&\left|\frac{1}{x-a}\int_a^x f(t)\,dt-\frac{1}{d-x}\int_x^d f(t)\,dt+t_n^{[a,x]}(x)-t_n^{[x,d]}(x)\right| \\
&\le\frac{B(\alpha+1,n-1)\,B(2,\alpha+n)}{(n-2)!}L\left[(x-a)^{\alpha+n}+(d-x)^{\alpha+n}\right].
\end{aligned}$$

Remark 5. *If we suppose $b=d$ in both cases $[c,d]\subset[a,b]$ and $[a,b]\cap[c,d]=[c,b]$ the analogues results in Corollary 3 and Corollary 4 coincides.*

References

[1] G. A. Anastassiou, Ostrowski type inequalities, Proc. Amer. Math. Soc., 123 (1995), pp. 3775-3781

[2] G. A. Anastassiou, Univariate Ostrowski inequalities, Monatshefte für Mathematik, 135 (2002), 175-189.

[3] N. S. Barnett, P. Cerone, S. S. Dragomir and A. M. Fink, Comparing two integral means for absolutely continuous mappings whose derivatives are in $L_\infty [a, b]$ and applications, Computers and Math.With Appl. 44 (2002), 241-251.

[4] M. Matić, J.Pečarić, Two-point Ostrowski inequality, Math. Inequal. Appl. 4(2) (2001), 215-221.

[5] G. V. Milovanović, On some integral inequalities, Univ. Beograd, Publ. Elektrotehn. Fak. Ser. Mat. Fiz., No 498-541, (1975), pp. 119-144

[6] G. V. Milovanović, O nekim funkcionalnim nejednakostima, Univ. Beograd, Publ. Elektrotehn. Fak. Ser. Mat. Fiz., No 599, (1977), pp. 1-59

[7] D. S. Mitrinović, J. E. Pečarić, and A. M. Fink, Inequalities for functions and their Integrals and Derivatives, Kluwer Academic Publishers, Dordrecht, 1994.

[8] J. Pečarić, On the Čebyšev inequality, Bul. Inst. Politehn. Timisoara 25 (39) (1980), 10-11.

JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS,VOL.5,NO.1,25-52,2007,COPYRIGHT 2007 EUDOXUS PRESS,LLC

Quantitative Estimates for Distance Between Fuzzy Wavelet Type Operators

by

George A. Anastassiou
Department of Mathematical Sciences
University of Memphis
Memphis, TN 38152 U.S.A.
E-mail address: ganastss@memphis.edu

2000 AMS Subject Classification: 26D15, 26E50, 26E99, 41A17, 41A25, 41A99, 47S40.

Key Words and Phrases: Fuzzy wavelet operator, fuzzy modulus of continuity, fuzzy inequalities, distances between fuzzy operators.

Abstract

The basic fuzzy wavelet type operators A_k, B_k, C_k, D_k, $k \in \mathbb{Z}$ were first introduced in [4], where they were studied among others for their pointwise/uniform convergence with rates to the fuzzy unit operator I. Here we continue this study by estimating the fuzzy distances between these operators. We give the pointwise convergence with rates of these distances to zero. The related approximation is of higher order since we involve these higher order fuzzy derivatives of the engaged fuzzy continuous function f. The derived Jackson type inequalities involve the fuzzy (first) modulus of continuity. Some comparison inequalities are also given so we get better upper bounds to the distances we study. The defining these operators scaling function φ is of compact support in $[-a, a]$, $a > 0$ and is not assumed to be orthogonal. Initially we estimate similarly the distances of $(B_k f)(x)$, $(C_k f)(x)$, $(D_k f)(x)$ from $f(x - \frac{a}{2^k})$, $x \in \mathbb{R}$. The main results of the paper rely on these initial results.

0. Introduction

This work is motivated by [4], especially from the following result:

Theorem 1 ([4]). *Let f be a function from $\mathbb{R}$ into the fuzzy real numbers $\mathbb{R}_{\mathcal{F}}$ which is fuzzy continuous. Let $\varphi(x)$ be a real valued bounded scaling function with* supp $\varphi(x) \subseteq [-a, a]$, *$a > 0$, $\varphi(x) \geq 0$, such that $\sum\limits_{j=-\infty}^{\infty} \varphi(x-j) \equiv 1$ on $\mathbb{R}$. For $k \in \mathbb{Z}$, $x \in \mathbb{R}$ put*

$$(B_k f)(x) := \text{fuzzy sum} \sum_{j=-\infty}^{\infty} f\left(\frac{j}{2^k}\right) \odot \varphi(2^k x - j)$$

($\odot$ denotes fuzzy multiplication). Clearly B_k is a fuzzy wavelet type operator.

Then the fuzzy distance

$$D((B_k f)(x), f(x)) \leq \omega_1^{(\mathcal{F})}\left(f, \frac{a}{2^k}\right),$$

any $x \in \mathbb{R}$, and $k \in \mathbb{Z}$, and

$$\sup_{x \in \mathbb{R}} D((B_k f)(x), f(x)) \leq \omega_1^{(\mathcal{F})}\left(f, \frac{a}{2^k}\right).$$

Here $\omega_1^{(\mathcal{F})}$ stands for the fuzzy (first) modulus of continuity. If f is fuzzy uniformly continuous then $\lim\limits_{k \to +\infty} B_k f = f$ uniformly with rates.

All fuzzy wavelet type operators A_k, B_k, C_k, D_k, $k \in \mathbb{Z}$ are reintroduced here and we find upper bounds to their distances $D((B_k f)(x), (C_k f)(x))$, $D((D_k f)(x), (C_k f)(x))$, $D((B_k f)(x), (D_k f)(x))$, $D((A_k f)(x), (B_k f)(x))$, $D((A_k f)(x), (C_k f)(x))$, and $D((A_k f)(x), (D_k f)(x))$. Their proofs rely a lot on the found here upper bounds for $D((B_k f)(x), f(x - \frac{a}{2^k}))$, $D((C_k f)(x), f(x - \frac{a}{2^k}))$ and $D((D_k f)(x), f(x - \frac{a}{2^k}))$, $k \in \mathbb{Z}$, $x \in \mathbb{R}$. The produced associated inequalities involve fuzzy (first) moduli of continuity of the engaged function and its fuzzy derivatives.

For fuzzy uniformly continuous functions f and its likewise derivatives we obtain pointwise convergence with rates to zero of all the above mentioned distances among the stated sequences of fuzzy wavelet type operators. For these see Section 2.

1. Background

We start with

Definition A (see [9]). Let $\mu: \mathbb{R} \to [0,1]$ with the following properties:

(i) is *normal*, i.e., $\exists x_0 \in \mathbb{R}$; $\mu(x_0) = 1$.

(ii) $\mu(\lambda x + (1-\lambda)y) \geq \min\{\mu(x), \mu(y)\}$, $\forall x, y \in \mathbb{R}$, $\forall \lambda \in [0,1]$ (μ is called a convex fuzzy subset).

(iii) μ is *upper semicontinuous* on $\mathbb{R}$, i.e., $\forall x_0 \in \mathbb{R}$ and $\forall \varepsilon > 0$, $\exists$ neighborhood $V(x_0)$: $\mu(x) \leq \mu(x_0) + \varepsilon$, $\forall x \in V(x_0)$.

(iv) The set $\overline{\operatorname{supp}(\mu)}$ is compact in $\mathbb{R}$ (where $\operatorname{supp}(\mu) := \{x \in \mathbb{R};\ \mu(x) > 0\}$).

We call μ a *fuzzy real number*. Denote the set of all μ with $\mathbb{R}_{\mathcal{F}}$.

E.g., $\mathcal{X}_{\{x_0\}} \in \mathbb{R}_{\mathcal{F}}$, for any $x_0 \in \mathbb{R}$, where $\mathcal{X}_{\{x_0\}}$ is the characteristic function at x_0. For $0 < r \leq 1$ and $\mu \in \mathbb{R}_{\mathcal{F}}$ define $[\mu]^r := \{x \in \mathbb{R}\colon\ \mu(x) \geq r\}$ and

$$[\mu]^0 := \overline{\{x \in \mathbb{R} : \mu(x) > 0\}}.$$

Then it is well known that for each $r \in [0,1]$, $[\mu]^r$ is a closed and bounded interval of $\mathbb{R}$. For $u, v \in \mathbb{R}_{\mathcal{F}}$ and $\lambda \in \mathbb{R}$, we define uniquely the sum $u \oplus v$ and the product $\lambda \odot u$ by

$$[u \oplus v]^r = [u]^r + [v]^r, \quad [\lambda \odot u]^r = \lambda [u]^r, \qquad \forall r \in [0,1],$$

where $[u]^r + [v]^r$ means the usual addition of two intervals (as subsets of $\mathbb{R}$) and $\lambda[u]^r$ means the usual product between a scalar and a subset of $\mathbb{R}$ (see, e.g., [9]). Notice $1 \odot u = u$ and it holds $u \oplus v = v \oplus u$, $\lambda \odot u = u \odot \lambda$. If $0 \leq r_1 \leq r_2 \leq 1$ then $[u]^{r_2} \subseteq [u]^{r_1}$. Actually $[u]^r = [u_-^{(r)}, u_+^{(r)}]$, where $u_-^{(r)} \leq u_+^{(r)}$, $u_-^{(r)}$, $u_+^{(r)} \in \mathbb{R}$, $\forall r \in [0,1]$.

Define

$$D: \mathbb{R}_{\mathcal{F}} \times \mathbb{R}_{\mathcal{F}} \to \mathbb{R}_+$$

by

$$D(u,v) := \sup_{r \in [0,1]} \max\{|u_-^{(r)} - v_-^{(r)}|, |u_+^{(r)} - v_+^{(r)}|\},$$

where $[v]^r = [v_-^{(r)}, v_+^{(r)}]$; $u, v \in \mathbb{R}_{\mathcal{F}}$. We have that D is a metric on $\mathbb{R}_{\mathcal{F}}$. Then $(\mathbb{R}_{\mathcal{F}}, D)$ is a complete metric space, see [9], with the properties

$$D(u \oplus w, v \oplus w) = D(u, v), \quad \forall u, v, w \in \mathbb{R}_{\mathcal{F}},$$

$$D(k \odot u, k \odot v) = |k| D(u, v), \quad \forall u, v \in \mathbb{R}_{\mathcal{F}}, \ \forall k \in \mathbb{R},$$

$$D(u \oplus v, w \oplus e) \leq D(u, w) + D(v, e), \quad \forall u, v, w, e \in \mathbb{R}_{\mathcal{F}}.$$

Let $f, g \colon \mathbb{R} \to \mathbb{R}_{\mathcal{F}}$ be *fuzzy real number valued functions*. The distance between f, g is defined by

$$D^*(f, g) := \sup_{z \in \mathbb{R}} D(f(x), g(x)).$$

We need

Lemma 4.1 ([5]). (i) *If we denote $\tilde{o} := \mathcal{X}_{\{0\}}$, then $\tilde{o} \in \mathbb{R}_{\mathcal{F}}$ is the neutral element with respect to $\oplus$, i.e., $u \oplus \tilde{o} = \tilde{o} \oplus u = u$, $\forall u \in \mathbb{R}_{\mathcal{F}}$,*

(ii) *With respect to $\tilde{o}$, none of $u \in \mathbb{R}_{\mathcal{F}}$, $u \neq \tilde{o}$ has opposite in $\mathbb{R}_{\mathcal{F}}$.*

(iii) *Let $a, b \in \mathbb{R} \colon a \cdot b \geq 0$, and any $u \in \mathbb{R}_{\mathcal{F}}$ we have $(a + b) \odot u = a \odot u \oplus b \odot v$.*

For general $a, b \in \mathbb{R}$, the above property is false.

(iv) *For any $\lambda \in \mathbb{R}$ and any $u, v \in \mathbb{R}_{\mathcal{F}}$, we have $\lambda \odot (u \oplus v) = \lambda \odot u \oplus \lambda \odot v$.*

(v) *For any $\lambda, \mu \in \mathbb{R}$ and $u \in \mathbb{R}_{\mathcal{F}}$, we have $\lambda \odot (\mu \odot u) = (\lambda \cdot \mu) \odot u$.*

(vi) *If we denote $\|u\|_{\mathcal{F}} := D(u, \tilde{o})$, $\forall u \in \mathbb{R}_{\mathcal{F}}$, then $\|\cdot\|_{\mathcal{F}}$ has the properties of a usual norm on $\mathbb{R}_{\mathcal{F}}$, i.e.,*

$$\|u\|_{\mathcal{F}} = 0 \ \text{ iff } u = \tilde{o}, \ \|\lambda \odot u\|_{\mathcal{F}} = |\lambda| \cdot \|u\|_{\mathcal{F}},$$

$$\|u \oplus v\|_{\mathcal{F}} \leq \|u\|_{\mathcal{F}} + \|v\|_{\mathcal{F}}, \ \|u\|_{\mathcal{F}} - \|v\|_{\mathcal{F}} \leq D(u, v).$$

Notice that $(\mathbb{R}_{\mathcal{F}}, \oplus, \odot)$ is *not* a linear space over $\mathbb{R}$, and consequently $(\mathbb{R}_{\mathcal{F}}, \|\cdot\|_{\mathcal{F}})$ is *not* a normed space.

Here $\sum^*$ stands for the fuzzy summation.

We use the following

Definition 1 ([3]). Let $f \colon \mathbb{R} \to \mathbb{R}_{\mathcal{F}}$ be a fuzzy real number valued function. We define the (*first*) *fuzzy modulus of continuity* of f by

$$\omega_1^{(\mathcal{F})}(f, \delta) := \sup_{\substack{x, y \in \mathbb{R} \\ |x - y| \leq \delta}} D(f(x), f(y)), \quad \delta > 0.$$

Definition 2 ([3]). Let $f:\mathbb{R}\to\mathbb{R}_{\mathcal{F}}$. If $D(f(x),\tilde{o})\leq M_1$, $\forall x\in\mathbb{R}$, $M_1>0$, we call f a *bounded fuzzy real number valued function.*

Definition 3 ([3]). Let $f:\mathbb{R}\to\mathbb{R}_{\mathcal{F}}$. We say that f is *fuzzy continuous at* $a\in\mathbb{R}$ if whenever $x_n\to a$, then $D(f(x_n),f(a))\to 0$. If f is continuous for every $a\in\mathbb{R}$, then we call f a *fuzzy continuous real number valued function.* We denote it as $f\in C(\mathbb{R},\mathbb{R}_{\mathcal{F}})$.

Definition 4 ([3]). Let $f:\mathbb{R}\to\mathbb{R}_{\mathcal{F}}$. We call f a *fuzzy uniformly continuous real number valued function*, iff for any $\varepsilon>0$ there exists $\delta>0$: whenver $|x-y|\leq\delta$, $x,y\in\mathbb{R}$, implies that $D(f(x),f(y))\leq\varepsilon$. We denote it as $f\in C^U_{\mathcal{F}}(\mathbb{R})$.

Proposition 1 ([3]). *Let* $f\in C^U_{\mathcal{F}}(\mathbb{R})$. *Then* $\omega_1^{(\mathcal{F})}(f,\delta)<+\infty$, *any* $\delta>0$.

Denote $f:\mathbb{R}\to\mathbb{R}_{\mathcal{F}}$ which is bounded and fuzzy continuous, as $f\in C_b(\mathbb{R},\mathbb{R}_{\mathcal{F}})$.

Proposition 2 ([3]). *It holds*

(i) $\omega_1^{(\mathcal{F})}(f,\delta)$ *is nonnegative and nondecreasing in* $\delta>0$, *any* $f:\mathbb{R}\to\mathbb{R}_{\mathcal{F}}$.

(ii) $\lim_{\delta\downarrow 0}\omega_1^{(\mathcal{F})}(f,\delta)=\omega_1^{(\mathcal{F})}(f,0)=0$, *iff* $f\in C^U_{\mathcal{F}}(\mathbb{R})$.

(iii) $\omega_1^{(\mathcal{F})}(f,\delta_1+\delta_2)\leq\omega_1^{(\mathcal{F})}(f,\delta_1)+\omega_1^{(\mathcal{F})}(f,\delta_2)$, $\delta_1,\delta_2>0$, *any* $f:\mathbb{R}\to\mathbb{R}_{\mathcal{F}}$.

(iv) $\omega_1^{(\mathcal{F})}(f,n\delta)\leq n\omega_1^{(\mathcal{F})}(f,\delta)$, $\delta>0$, $n\in\mathbb{N}$, *any* $f:\mathbb{R}\to\mathbb{R}_{\mathcal{F}}$.

(v) $\omega_1^{(\mathcal{F})}(f,\lambda\delta)\leq\lceil\lambda\rceil\omega_1^{(\mathcal{F})}(f,\delta)\leq(\lambda+1)\omega_1^{(\mathcal{F})}(f,\delta)$, $\lambda>0$, $\delta>0$, *where* $\lceil\cdot\rceil$ *is the ceiling of the number, any* $f:\mathbb{R}\to\mathbb{R}_{\mathcal{F}}$.

(vi) $\omega_1^{(\mathcal{F})}(f\oplus g,\delta)\leq\omega_1^{(\mathcal{F})}(f,\delta)+\omega_1^{(\mathcal{F})}(g,\delta)$, $\delta>0$, *any* $f,g:\mathbb{R}\to\mathbb{R}_{\mathcal{F}}$.

(vii) $\omega_1^{(\mathcal{F})}(f,\cdot)$ *is continuous on* $\mathbb{R}_+$, *for* $f\in C^U_{\mathcal{F}}(\mathbb{R})$.

We also use

Definition 5 (see [9]). Let $x,y\in\mathbb{R}_{\mathcal{F}}$. If there exists a $z\in\mathbb{R}_{\mathcal{F}}$ such that $x=y\oplus z$,then we call z the *H-difference* of x and y. Denoted by $z:=x-y$.

Definition 3.3 (see [9]). Let $T:=[x_0,x_0+\beta]\subset\mathbb{R}$, with $\beta>0$. A function $f:T\to\mathbb{R}_{\mathcal{F}}$ is *fuzzy differentiable* at $x\in T$ if there exists a $f'(x)\in\mathbb{R}_{\mathcal{F}}$ such that the limits in D-metric

$$\lim_{h\to 0^+}\frac{f(x+h)-f(x)}{h},\quad \lim_{h\to 0^+}\frac{f(x)-f(x-h)}{h}$$

exist and are equal to $f'(x)$. We call f' the *fuzzy derivative* of f at x. If f is fuzzy differentiable at any $x \in T$, we call f *fuzzy differentiable* and it has *fuzzy derivative over* T the function f'.

We need also a particular case of the *fuzzy Henstock integral* ($\delta(x) = \frac{\delta}{2}$) introduced in [9], Definition 2.1. That is,

Definition 13.14 (see [7], p. 644). Let $f\colon [a,b] \to \mathbb{R}_{\mathcal{F}}$. We say that f is fuzzy-Riemann integrable to $I \in \mathbb{R}_{\mathcal{F}}$ if for any $\varepsilon > 0$, there exists $\delta > 0$ such that for any division $P = \{[u,v],\xi\}$ of $[a,b]$ with the norm $\Delta(P) < \delta$, we have

$$D\left(\sum_{P}^{*}(v-u)\odot f(\xi), I\right) < \varepsilon.$$

We prefer to write

$$I := (FR)\int_a^b f(x)dx.$$

We also call an f as above (FR)-*integrable*.

Corollary 13.2 ([7], p. 644). *If f is fuzzy continuous from $[a,b]$ into $\mathbb{R}_{\mathcal{F}}$ then f is (FR)-integrable on $[a,b]$.*

Theorem 3.4 ([8]). *If $f, g\colon [c,d] \to \mathbb{R}_{\mathcal{F}}$ are (FR)-integrable fuzzy functions, and α, β are real numbers, then*

$$\begin{aligned}(FR)\int_c^d (\alpha f(x) \oplus \beta g(x))dx = \alpha(FR)\int_c^d f(x)dx \\ \oplus \beta(FR)\int_c^d g(x)dx.\end{aligned}$$

We need also

Lemma 1 ([3]). *If $f, g\colon [a,b] \subseteq \mathbb{R} \to \mathbb{R}_{\mathcal{F}}$ are fuzzy continuous, then the function $F\colon [a,b] \to \mathbb{R}_+$ defined by $F(x) := D(f(x), g(x))$ is continuous on $[a,b]$ and*

$$D\left((FR)\int_a^b f(u)du, (FR)\int_a^b g(u)du\right) \le \int_a^b F(x)dx.$$

Lemma 3 ([3]). *Let $f\colon [a,b] \subset \mathbb{R} \to \mathbb{R}_{\mathcal{F}}$ be fuzzy continuous. Then $(FR)\int_a^x f(t)dt$ is a fuzzy continuous function in $x \in [a,b]$.*

Lemma 4 ([3]). *Let $f \in C(\mathbb{R}, \mathbb{R}_{\mathcal{F}})$, $r \in \mathbb{N}$. Then the following integrals*

$$(FR)\int_a^{s_{r-1}} f(s_r)ds_r, (FR)\int_a^{s_{r-2}}\left(\int_a^{s_{r-1}} f(s_r)ds_r\right)ds_{r-1}, \ldots,$$

$$(FR)\left(\int_a^s \int_a^{s_1} \cdots \left(\int_a^{s_{r-2}}\left(\int_a^{s_{r-1}} f(s_r)ds_r\right)ds_{r-1}\right)\cdots\right)ds_1,$$

are fuzzy continuous functions in s_{r-1}, $s_{r-2}, \ldots, s$, respectively. Here s_{r-1}, $s_{r-2}, \ldots, s \geq a$ and all are real numbers.

Here we use a lot the following fuzzy Taylor's formula.

Theorem 1 ([3]). *Let $T := [x_0, x_0 + \beta] \subset \mathbb{R}$, with $\beta > 0$. We assume that $f^{(i)}: T \to \mathbb{R}_{\mathcal{F}}$ are fuzzy differentiable for all $i = 0, 1, \ldots, n-1$, for any $x \in T$ (i.e., there exist in $\mathbb{R}_{\mathcal{F}}$ the H-differences $f^{(i)}(x+h) - f^{(i)}(x)$, $f^{(i)}(x) - f^{(i)}(x-h)$, $i = 0, 1, \ldots, n-1$ for all small $0 < h < \beta$. Furthermore there exist $f^{(i+1)}(x) \in \mathbb{R}_{\mathcal{F}}$ such that the limits in D-distance exist and*

$$f^{(i+1)}(x) = \lim_{h \to 0^+} \frac{f^{(i)}(x+h) - f^{(i)}(x)}{h} = \lim_{h \to 0^+} \frac{f^{(i)}(x) - f^{(i)}(x-h)}{h},$$

for all $i = 0, 1, \ldots, n-1$). Also we assume that $f^{(i)}$, $i = 0, 1, \ldots, n$ are fuzzy continuous on T. Then for $s \geq a$, $s, a \in T$ we obtain

$$f(s) = f(a) \oplus f'(a) \odot (s-a) \oplus f''(a) \odot \frac{(s-a)^2}{2!} \oplus \cdots \oplus f^{(n-1)}(a) \odot \frac{(s-a)^{n-1}}{(n-1)!} \oplus R_n(a,s)$$

where

$$R_n(a,s) := (FR)\int_a^s \left(\int_a^{s_1} \cdots \left(\int_a^{s_{n-1}} f^{(n)}ds_n\right)ds_{n-1}\right)\cdots\right)ds_1.$$

Here $R_n(a,s)$ is fuzzy continuous over T as a function of s.

Note 1. This formula is invalid when $s < a$, as it is based on Theorem 3.6 of [9].

We denote by $C^N(\mathbb{R}, \mathbb{R}_{\mathcal{F}})$ the space of N-times continuously differentiable in the fuzzy sense functions from $\mathbb{R}$ into $\mathbb{R}_{\mathcal{F}}$, $N \in \mathbb{N}$. We also denote by $C_{\mathcal{F}}^{NU}(\mathbb{R})$, $N \in \mathbb{N}$, the space of functions $f: \mathbb{R} \to \mathbb{R}_{\mathcal{F}}$, such that the fuzzy derivatives exist up to order N and all f, $f', \ldots, f^{(N)}$ are fuzzy uniformly continuous.

Finally we make use of

Lemma 2 ([4]). *Let $f: \mathbb{R} \to \mathbb{R}_{\mathcal{F}}$ be fuzzy continuous and bounded function, i.e. $\exists M_1 > 0$: $D(f(x), \tilde{o}) \leq M_1$, $\forall x \in \mathbb{R}$. Let also g: from interval $J \subseteq \mathbb{R} \to \mathbb{R}_+$ continuous and bounded function. Then $f(x) \odot g(x)$ is a fuzzy continuous function on J.*

2. Results

We present our first main result.

Theorem 2. *Let $f \in C^N(\mathbb{R}, \mathbb{R}_{\mathcal{F}})$, $N \in \mathbb{N}$ and the scaling function $\varphi(x)$ a real valued bounded function with* supp $\varphi(x) \subseteq [-a, a]$, $0 < a < +\infty$, $\varphi(x) \geq 0$, *such that* $\sum_{j=-\infty}^{\infty} \varphi(x - j) \equiv 1$ *on* $\mathbb{R}$. *For* $k \in \mathbb{Z}$, $x \in \mathbb{R}$ *put*

$$(B_k f)(x) := \sum_{j=-\infty}^{\infty}{}^{*} f\left(\frac{j}{2^k}\right) \odot \varphi(2^k x - j), \tag{1}$$

which is a fuzzy wavelet type operator. Then it holds

$$\begin{aligned} D\left((B_k f)(x), f\left(x - \frac{a}{2^k}\right)\right) &\leq \sum_{i=1}^{N-1} \frac{a^i}{2^{i(k-1)} i!} \left(D(f^{(i)}(x), \tilde{o}) + \omega_1^{(\mathcal{F})}\left(f^{(i)}, \frac{a}{2^k}\right)\right) \\ &\quad + \frac{a^N}{2^{N(k-1)} N!} \left(D(f^{(N)}(x), \tilde{o}) + 3\omega_1^{(\mathcal{F})}\left(f^{(N)}, \frac{a}{2^k}\right)\right) \\ &=: \beta_k\left(\frac{a}{2^k}\right), \quad \text{for any } x \in \mathbb{R},\ k \in \mathbb{Z}. \end{aligned} \tag{2}$$

If $f \in C_{\mathcal{F}}^{NU}(\mathbb{R})$ *and as* $k \to +\infty$ *we obtain with rates that*

$$\lim_{k \to +\infty} D\left((B_k f)(x), f\left(x - \frac{a}{2^k}\right)\right) = 0.$$

Corollary 1 (to Theorem 2, for $N = 1$). *It holds*

$$D\left(B_k f)(x), f\left(x - \frac{a}{2^k}\right)\right) \leq \frac{a}{2^{k-1}} \left(D(f'(x), \tilde{o}) + 3\omega_1^{(\mathcal{F})}\left(f', \frac{a}{2^k}\right)\right),$$

for any $x \in \mathbb{R}$, $k \in \mathbb{Z}$.

Corollary 2 (to Theorem 2). *The following improvement of (2) holds*

$$D\left((B_k f)(x), f\left(x - \frac{a}{2^k}\right)\right) \leq \min\left(2\omega_1^{(\mathcal{F})}\left(f, \frac{a}{2^k}\right), \beta_k\left(\frac{a}{2^k}\right)\right),$$

for any $x \in \mathbb{R}$, $k \in \mathbb{Z}$.

Proof. From Theorem 1, inequality (2), of [4] we get

$$D((B_k f)(x), f(x)) \leq \omega_1^{(\mathcal{F})}\left(f, \frac{a}{2^k}\right).$$

Then we observe

$$D\left((B_kf)(x), f\left(x - \frac{a}{2^k}\right)\right) \le D((B_kf)(x), f(x)) + D\left(f(x), f\left(x - \frac{a}{2^k}\right)\right)$$
$$\le 2\omega_1^{(\mathcal{F})}\left(f, \frac{a}{2^k}\right). \qquad \square$$

Proof of Theorem 2. Because φ is of compact support in $[-a, a]$ we see that

$$(B_kf)(x) = \sum_{\substack{j \\ 2^kx - j \in [-a,a]}}^{*} f\left(\frac{j}{2^k}\right) \odot \varphi(2^kx - j). \tag{3}$$

That is for specific $x \in \mathbb{R}$, $k \in \mathbb{Z}$ we have the j's in (3) to fulfill

$$x - \frac{a}{2^k} \le \frac{j}{2^k} \le x + \frac{a}{2^k}.$$

Using the fuzzy Taylor formula (Theorem 1, [3]) we get

$$f\left(\frac{j}{2^k}\right) = \sum_{i=0}^{N-1}{}^{*} f^{(i)}\left(x - \frac{a}{2^k}\right) \odot \frac{\left(\frac{j+a}{2^k} - x\right)^i}{i!} \oplus \mathcal{R}_N\left(x - \frac{a}{2^k}, \frac{j}{2^k}\right), \tag{4}$$

where

$$\mathcal{R}_N\left(x - \frac{a}{2^k}, \frac{j}{2^k}\right) := (FR)\int_{x-\frac{a}{2^k}}^{\frac{j}{2^k}}\left(\int_{x-\frac{a}{2^k}}^{s_1}\left(\right.\right.$$
$$\left.\left.\cdots\left(\int_{x-\frac{a}{2^k}}^{s_{N-1}} f^{(N)}(s_N)ds_N\right)ds_{N-1}\cdots\right)ds_1. \tag{5}$$

Then

$$f\left(\frac{j}{2^k}\right) \odot \varphi(2^kx - j) = \sum_{i=0}^{N-1}{}^{*} f^{(i)}\left(x - \frac{a}{2^k}\right) \odot \frac{\left(\frac{j+a}{2^k} - x\right)^i}{i!} \odot \varphi(2^kx - j)$$
$$\oplus \mathcal{R}_N\left(x - \frac{a}{2^k}, \frac{j}{2^k}\right) \odot \varphi(2^kx - j)$$

and

$$\sum_{\substack{j \\ 2^kx - j \in [-a,a]}}^{*} f\left(\frac{j}{2^k}\right) \odot \varphi(2^kx - j) = \sum_{\substack{j \\ 2^kx - j \in [-a,a]}}^{*} \sum_{i=0}^{N-1}{}^{*} f^{(i)}\left(x - \frac{a}{2^k}\right) \odot \frac{\left(\frac{j+a}{2^k} - x\right)^i}{i!}$$
$$\odot \varphi(2^kx - j) \oplus \sum_{\substack{j \\ 2^kx - j \in [-a,a]}}^{*} \mathcal{R}_N\left(x - \frac{a}{2^k}, \frac{j}{2^k}\right) \odot \varphi(2^kx - j).$$

That is we have

$$(B_k f)(x) = \sum_{i=0}^{N-1}{}^{*} \sum_{\substack{j \\ 2^k x - j \in [-a,a]}}^{*} f^{(i)}\left(x - \frac{a}{2^k}\right) \odot \frac{\left(\frac{j+a}{2^k} - x\right)^i}{i!} \odot \varphi(2^k x - j)$$

$$\oplus \sum_{\substack{j \\ 2^k x - j \in [-a,a]}}^{*} \mathcal{R}_N\left(x - \frac{a}{2^k}, \frac{j}{2^k}\right) \odot \varphi(2^k x - j). \tag{6}$$

Next we estimate

$$D\left((B_k f)(x), f\left(x - \frac{a}{2^k}\right)\right)$$

$$= D\left((B_k f)(x), \sum_{\substack{j \\ 2^k x - j \in [-a,a]}}^{*} \varphi(2^k x - j) \odot f\left(x - \frac{a}{2^k}\right)\right)$$

$$= D\Bigg(\sum_{i=1}^{N-1}{}^{*} \sum_{\substack{j \\ 2^k x - j \in [-a,a]}}^{*} f^{(i)}\left(x - \frac{a}{2^k}\right) \odot \frac{\left(\frac{j+a}{2^k} - x\right)^i}{i!} \odot \varphi(2^k x - j)$$

$$\oplus \sum_{\substack{j \\ 2^k x - j \in [-a,a]}}^{*} \mathcal{R}_N\left(x - \frac{a}{2^k}, \frac{j}{2^k}\right) \odot \varphi(2^k x - j), \tilde{o}\Bigg)$$

$$\leq \sum_{i=1}^{N-1} \sum_{\substack{j \\ 2^k x - j \in [-a,a]}} \frac{\left(\frac{j+a}{2^k} - x\right)^i}{i!} \varphi(2^k x - j) D\left(f^{(i)}\left(x - \frac{a}{2^k}\right), \tilde{o}\right)$$

$$+ \sum_{\substack{j \\ 2^k x - j \in [-a,a]}} \varphi(2^k x - j) D\left(\mathcal{R}_N\left(x - \frac{a}{2^k}, \frac{j}{2^k}\right), \tilde{o}\right)$$

$$\leq \sum_{i=1}^{N-1} \sum_{\substack{j \\ 2^k x - j \in [-a,a]}} \frac{a^i}{2^{i(k-1)} i!} \varphi(2^k x - j) \left(D(f^{(i)}(x), \tilde{o}) + \omega_1^{(\mathcal{F})}\left(f^{(i)}, \frac{a}{2^k}\right)\right)$$

$$+ \sum_{\substack{j \\ 2^k x - j \in [-a,a]}} \varphi(2^k x - j) D\left(\mathcal{R}_N\left(x - \frac{a}{2^k}, \frac{j}{2^k}\right), \tilde{o}\right) =: (*).$$

Next we work on

$$D\left(\mathcal{R}_N\left(x - \frac{a}{2^k}, \frac{j}{2^k}\right), \tilde{o}\right)$$

$$= D\Bigg(\mathcal{R}_N\left(x - \frac{a}{2^k}, \frac{j}{2^k}\right) \oplus f^{(N)}\left(x - \frac{a}{2^k}\right)$$

$$\odot \frac{\left(\frac{j+a}{2^k} - x\right)^N}{N!}, f^{(N)}\left(x - \frac{a}{2^k}\right) \odot \frac{\left(\frac{j+a}{2^k} - x\right)^N}{N!}\Bigg)$$

$$\leq D\left(\mathcal{R}_N\left(x-\frac{a}{2^k},\frac{j}{2^k}\right),f^{(N)}\left(x-\frac{a}{2^k}\right)\odot\frac{\left(\frac{j+a}{2^k}-x\right)^N}{N!}\right)$$
$$+D\left(f^{(N)}\left(x-\frac{a}{2^k}\right)\odot\frac{\left(\frac{j+a}{2^k}-x\right)^N}{N!},\tilde{o}\right).$$

So that

$$D\left(\mathcal{R}_N\left(x-\frac{a}{2^k},\frac{j}{2^k}\right),\tilde{o}\right)\leq D\left(\mathcal{R}_N\left(x-\frac{a}{2^k},\frac{j}{2^k}\right),f^{(N)}\left(x-\frac{a}{2^k}\right)\odot\frac{\left(\frac{j+a}{2^k}-x\right)^N}{N!}\right)$$
$$+\frac{a^N}{2^{N(k-1)}N!}\left(D(f^{(N)}(x),\tilde{o})+\omega_1^{(\mathcal{F})}\left(f^{(N)},\frac{a}{2^k}\right)\right). \tag{7}$$

At the end we estimate

$$D\left(\mathcal{R}_N\left(x-\frac{a}{2^k},\frac{j}{2^k}\right),f^{(N)}\left(x-\frac{a}{2^k}\right)\odot\frac{\left(\frac{j+a}{2^k}-x\right)^N}{N!}\right)$$
$$=D\Bigg((FR)\int_{x-\frac{a}{2^k}}^{\frac{j}{2^k}}\left(\int_{x-\frac{a}{2^k}}^{s_1}\left(\int_{x-\frac{a}{2^k}}^{s_{N-1}}f^{(N)}(s_N)ds_N\right)ds_{N-1}\cdots\right)ds_1,f^{(N)}\left(x-\frac{a}{2^k}\right)$$
$$\odot\int_{x-\frac{a}{2^k}}^{\frac{j}{2^k}}\left(\int_{x-\frac{a}{2^k}}^{s_1}\cdots\left(\int_{x-\frac{a}{2^k}}^{s_{N-1}}1\,ds_N\right)ds_{N-1}\cdots\right)ds_1\Bigg)$$
$$=D\Bigg((FR)\int_{x-\frac{a}{2^k}}^{\frac{j}{2^k}}\left(\int_{x-\frac{a}{2^k}}^{s_1}\cdots\left(\int_{x-\frac{a}{2^k}}^{s_{N-1}}f^{(N)}(s_N)ds_N\right)ds_{N-1}\cdots\right)ds_1,$$
$$(FR)\int_{x-\frac{a}{2^k}}^{\frac{j}{2^k}}\left(\int_{x-\frac{a}{2^k}}^{s_1}\cdots\left(\int_{x-\frac{a}{2^k}}^{s_{N-1}}f^{(N)}\left(x-\frac{a}{2^k}\right)ds_N\right)ds_{N-1}\cdots\right)ds_1\Bigg)$$
$$\overset{\text{(by Lemma 1,4 of [3])}}{\leq}\int_{x-\frac{a}{2^k}}^{\frac{j}{2^k}}\left(\int_{x-\frac{a}{2^k}}^{s_1}\right.$$
$$\left.\cdots\left(\int_{x-\frac{a}{2^k}}^{s_{N-1}}D\left(f^{(N)}(s_N),f^{(N)}\left(x-\frac{a}{2^k}\right)\right)ds_N\right)ds_{N-1}\cdots\right)ds_1$$
$$\leq\int_{x-\frac{a}{2^k}}^{\frac{j}{2^k}}\left(\int_{x-\frac{a}{2^k}}^{s_1}\cdots\left(\int_{x-\frac{a}{2^k}}^{s_{N-1}}\omega_1^{(\mathcal{F})}\left(f^{(N)},\left|s_N+\frac{a}{2^k}-x\right|ds_N\right)ds_{N-1}\cdots\right)ds_1$$
$$\leq\omega_1^{(\mathcal{F})}\left(f^{(N)},\left|\frac{j}{2^k}+\frac{a}{2^k}-x\right|\right)\int_{x-\frac{a}{2^k}}^{\frac{j}{2^k}}\left(\int_{x-\frac{a}{2^k}}^{s_1}\cdots\left(\int_{x-\frac{a}{2^k}}^{s_{N-1}}1\,ds_N\right)ds_{N-1}\cdots\right)ds_1$$
$$=\omega_1^{(\mathcal{F})}\left(f^{(N)},\left|\frac{j}{2^k}+\frac{a}{2^k}-x\right|\right)\frac{\left(\frac{j}{2^k}-\left(x-\frac{a}{2^k}\right)\right)^N}{N!}\leq\frac{a^N}{2^{N(k-1)}N!}\omega_1^{(\mathcal{F})}\left(f^{(N)},\frac{a}{2^{k-1}}\right).$$

I.e. we have found that

$$D\left(\mathcal{R}_N\left(x-\frac{a}{2^k},\frac{j}{2^k}\right),f^{(N)}\left(x-\frac{a}{2^k}\right)\odot\frac{\left(\frac{j+a}{2^k}-x\right)^N}{N!}\right)$$
$$\leq\frac{a^N}{2^{N(k-1)}N!}\omega_1^{(\mathcal{F})}\left(f^{(N)},\frac{a}{2^{k-1}}\right). \tag{8}$$

Therefore by (7) and (8) we obtain

$$D\left(\mathcal{R}_N\left(x-\frac{a}{2^k},\frac{j}{2^k}\right),\tilde{o}\right)\leq\frac{a^N}{2^{N(k-1)}N!}\left(D(f^{(N)},(x),\tilde{o})+3\omega_1^{(\mathcal{F})}\left(f^{(N)},\frac{a}{2^k}\right)\right). \tag{9}$$

Using (9) into $(*)$ we get

$$\begin{aligned}(*)\leq&\sum_{i=1}^{N-1}\frac{a^i}{2^{i(k-1)}i!}\left(D(f^{(i)}(x),\tilde{o})+\omega_1^{(\mathcal{F})}\left(f^{(i)},\frac{a}{2^k}\right)\right)\\&+\frac{a^N}{2^{N(k-1)}N!}\left(D(f^{(N)}(x),\tilde{o})+3\omega_1^{(\mathcal{F})}\left(f^{(N)},\frac{a}{2^k}\right)\right).\end{aligned}$$

Inequality (2) has been established. □

The next related main result is given.

Theorem 3. *All assumptions are as in Theorem 2. Define for $k\in\mathbb{Z}$, $x\in\mathbb{R}$ the fuzzy wavelet type operator*

$$(D_kf)(x):=\sum_{j=-\infty}^{\infty}{}^{*}\,\delta_{kj}(f)\odot\varphi(2^kx-j), \tag{10}$$

where

$$\delta_{kj}(f):=\sum_{\tilde{r}=0}^{n}{}^{*}\,w_{\tilde{r}}\odot f\left(\frac{j}{2^k}+\frac{\tilde{r}}{2^kn}\right),\ n\in\mathbb{N},\ w_{\tilde{r}}\geq 0,\ \sum_{\tilde{r}=0}^{n}w_{\tilde{r}}=1. \tag{11}$$

Then it holds

$$\begin{aligned}&D\left((D_kf)(x),f\left(x-\frac{a}{2^k}\right)\right)\\&\leq\sum_{i=1}^{N-1}\frac{(2a+1)^i}{2^{ki}i!}\left(D(f^{(i)}(x),\tilde{o})+\omega_1^{(\mathcal{F})}\left(f^{(i)},\frac{a}{2^k}\right)\right)\\&\quad+\frac{(2a+1)^N}{2^{kN}N!}\left(D(f^{(N)}(x),\tilde{o})+3\omega_1^{(\mathcal{F})}\left(f^{(N)},\frac{a}{2^k}\right)+\omega_1^{(\mathcal{F})}\left(f^{(N)},\frac{1}{2^k}\right)\right)\\&=:\ \Delta_k\left(\frac{a}{2^k}\right),\quad \text{for any } x\in\mathbb{R},\ k\in\mathbb{Z}.\end{aligned} \tag{12}$$

If $f\in C_{\mathcal{F}}^{NU}(\mathbb{R})$, as $k\to\infty$, we obtain with rates that

$$\lim_{k\to+\infty}D\left((D_kf)(x),f\left(x-\frac{a}{2^k}\right)\right)=0.$$

Corollary 3 (to Theorem 3, for $N=1$). *It holds*

$$D\left((D_kf)(x),f\left(x-\frac{a}{2^k}\right)\right)\leq\left(\frac{a}{2^{k-1}}+\frac{1}{2^k}\right)\left(D(f'(x),\tilde{o})+3\omega_1^{(\mathcal{F})}\left(f',\frac{a}{2^k}\right)+\omega_1^{(\mathcal{F})}\left(f',\frac{1}{2^k}\right)\right),$$

for any $x\in\mathbb{R}$, $k\in\mathbb{Z}$.

Corollary 4 (to Theorem 3). *The following improvement of (12) holds*

$$D\left((D_kf)(x), f\left(x-\frac{a}{2^k}\right)\right) \leq \min\left(\left(\omega_1^{(\mathcal{F})}\left(f, \frac{a+1}{2^k}\right)+\omega_1^{(\mathcal{F})}\left(f, \frac{a}{2^k}\right)\right), \Delta_k\left(\frac{a}{2^k}\right)\right),$$

for any $x \in \mathbb{R}$, $k \in \mathbb{Z}$.

Proof. From Theorem 4, inequality (14) of [4] we obtain

$$D((D_kf)(x), f(x)) \leq \omega_1^{(\mathcal{F})}\left(f, \frac{a+1}{2^k}\right).$$

Then we see that

$$\begin{aligned} D\left((D_kf)(x), f\left(x-\frac{a}{2^k}\right)\right) &\leq D((D_kf)(x), f(x)) + D\left(f(x), f\left(x-\frac{a}{2^k}\right)\right) \\ &\leq \omega_1^{(\mathcal{F})}\left(f, \frac{a+1}{2^k}\right)+\omega_1^{(\mathcal{F})}\left(f, \frac{a}{2^k}\right). \qquad \square \end{aligned}$$

Proof of Theorem 3. Because φ is of compact support in $[-a, a]$ we observe that

$$(D_kf)(x) = \sum^{*}_{\substack{j \\ 2^kx-j\in[-a,a]}} \left(\sum^{n\,*}_{\tilde{r}=0} w_{\tilde{r}} \odot f\left(\frac{j}{2^k}+\frac{\tilde{r}}{2^kn}\right)\right) \odot \varphi(2^kx-j). \tag{13}$$

Again for specific $x \in \mathbb{R}$, $k \in \mathbb{Z}$ we have that the j's in (12) satisfy

$$x-\frac{a}{2^k} \leq \frac{j}{2^k} \leq x+\frac{a}{2^k}.$$

Using the fuzzy Taylor formula (Theorem 1, [3]) we get

$$\begin{aligned} f\left(\frac{j}{2^k}+\frac{\tilde{r}}{2^kn}\right) = &\sum^{N-1\,*}_{i=0} f^{(i)}\left(x-\frac{a}{2^k}\right) \odot \frac{\left(\frac{(j+a)}{2^k}+\frac{\tilde{r}}{2^kn}-x\right)^i}{i!} \\ &\oplus \mathcal{R}_N\left(x-\frac{a}{2^k}, \frac{j}{2^k}+\frac{\tilde{r}}{2^kn}\right), \end{aligned} \tag{14}$$

where

$$\begin{aligned} &\mathcal{R}_N\left(x-\frac{a}{2^k}, \frac{j}{2^k}+\frac{\tilde{r}}{2^kn}\right) \\ &:= (FR)\int_{x-\frac{a}{2^k}}^{\left(\frac{j}{2^k}+\frac{\tilde{r}}{2^kn}\right)}\left(\int_{x-\frac{a}{2^k}}^{s_1}\cdots\left(\int_{x-\frac{a}{2^k}}^{s_{N-1}} f^{(N)}(s_N)ds_N\right)ds_{N-1}\cdots\right)ds_1. \end{aligned} \tag{15}$$

Then

$$\begin{aligned} \delta_{kj}(f) = \sum^{n\,*}_{\tilde{r}=0} w_{\tilde{r}} \odot f\left(\frac{j}{2^k}+\frac{\tilde{r}}{2^kn}\right) = &\sum^{N-1\,*}_{i=0} f^{(i)}\left(x-\frac{a}{2^k}\right) \odot \sum_{\tilde{r}=0}^{n} w_{\tilde{r}} \frac{\left(\frac{(j+a)}{2^k}+\frac{\tilde{r}}{2^kn}-x\right)^i}{i!} \\ \oplus &\sum^{n\,*}_{\tilde{r}=0} w_{\tilde{r}} \odot \mathcal{R}_N\left(x-\frac{a}{2^k}, \frac{j}{2^k}+\frac{\tilde{r}}{2^kn}\right), \end{aligned} \tag{16}$$

and

$$(D_k f)(x) = \sum_{i=0}^{N-1}{}^{*} f^{(i)}\left(x - \frac{a}{2^k}\right) \odot \sum_{\substack{j \\ 2^k x - j \in [-a,a]}} \varphi(2^k x - j)\left(\sum_{\tilde{r}=0}^{n} w_{\tilde{r}} \frac{\left(\frac{(j+a)}{2^k} + \frac{\tilde{r}}{2^k n} - x\right)^i}{i!}\right)$$
$$\oplus \sum_{\substack{j \\ 2^k x - j \in [-a,a]}} \varphi(2^k x - j)\left(\sum_{\tilde{r}=0}^{n}{}^{*} w_{\tilde{r}} \odot \mathcal{R}_N\left(x - \frac{a}{2^k}, \frac{j}{2^k} + \frac{\tilde{r}}{2^k n}\right)\right). \tag{17}$$

Next we estimate

$$D\left((D_k f)(x), f\left(x - \frac{a}{2^k}\right)\right)$$
$$= D\left((D_k f)(x), \sum_{\substack{j \\ 2^k x - j \in [-a,a]}}{}^{*} \varphi(2^k x - j) \odot f\left(x - \frac{a}{2^k}\right)\right)$$
$$= D\left(\sum_{i=1}^{N-1}{}^{*} f^{(i)}\left(x - \frac{a}{2^k}\right) \odot \sum_{\substack{j \\ 2^k x - j \in [-a,a]}} \varphi(2^k x - j)\left(\sum_{\tilde{r}=0}^{n} w_{\tilde{r}} \frac{\left(\frac{(j+a)}{2^k} + \frac{\tilde{r}}{2^k n} - x\right)^i}{i!}\right)\right.$$
$$\left.\oplus \sum_{\substack{j \\ 2^k x - j \in [-a,a]}} \varphi(2^k x - j)\left(\sum_{\tilde{r}=0}^{n}{}^{*} w_{\tilde{r}} \odot \mathcal{R}_N\left(x - \frac{a}{2^k}, \frac{j}{2^k} + \frac{\tilde{r}}{2^k n}\right)\right), \tilde{o}\right)$$
$$\leq \sum_{i=1}^{N-1} \sum_{\substack{j \\ 2^k x - j \in [-a,a]}} \varphi(2^k x - j)\left(\sum_{\tilde{r}=0}^{n} w_{\tilde{r}} \frac{\left(\frac{(j+a)}{2^k} + \frac{\tilde{r}}{2^k n} - x\right)^i}{i!}\right) D\left(f^{(i)}\left(x - \frac{a}{2^k}\right), \tilde{o}\right)$$
$$+ \sum_{\substack{j \\ 2^k x - j \in [-a,a]}} \varphi(2^k x - j)\left(\sum_{\tilde{r}=0}^{n} w_{\tilde{r}} D\left(\mathcal{R}_N\left(x - \frac{a}{2^k}, \frac{j}{2^k} + \frac{\tilde{r}}{2^k n}\right), \tilde{o}\right)\right)$$
$$\leq \sum_{i=1}^{N-1} \frac{(2a+1)^i}{2^{ki} i!}\left(D(f^{(i)}(x), \tilde{o}) + \omega_1^{(\mathcal{F})}\left(f^{(i)}, \frac{a}{2^k}\right)\right)$$
$$+ \sum_{\substack{j \\ 2^k x - j \in [-a,a]}} \varphi(2^k x - j)\left(\sum_{\tilde{r}=0}^{n} w_{\tilde{r}} D\left(\mathcal{R}_N\left(x - \frac{a}{2^k}, \frac{j}{2^k} + \frac{\tilde{r}}{2^k n}\right), \tilde{o}\right)\right) =: (*).$$

Next we work on

$$D\left(\mathcal{R}_N\left(x - \frac{a}{2^k}, \frac{j}{2^k} + \frac{\tilde{r}}{2^k n}\right), \tilde{o}\right)$$
$$\leq D\left(\mathcal{R}_N\left(x - \frac{a}{2^k}, \frac{j}{2^k} + \frac{\tilde{r}}{2^k n}\right), f^{(N)}\left(x - \frac{a}{2^k}\right) \odot \frac{\left(\frac{(j+a)}{2^k} + \frac{\tilde{r}}{2^k n} - x\right)^N}{N!}\right)$$

$$+ D\left(f^{(N)}\left(x - \frac{a}{2^k}\right) \odot \frac{(\frac{(j+a)}{2^k} + \frac{\tilde{r}}{2^k n} - x)^N}{N!}, \tilde{o}\right).$$

So that

$$\begin{aligned}
&D\left(\mathcal{R}_N\left(x - \frac{a}{2^k}, \frac{j}{2^k} + \frac{\tilde{r}}{2^k n}\right), \tilde{o}\right) \\
&\leq D\left(\mathcal{R}_N\left(x - \frac{a}{2^k}, \frac{j}{2^k} + \frac{\tilde{r}}{2^k n}\right), f^{(N)}\left(x - \frac{a}{2^k}\right) \odot \frac{(\frac{(j+a)}{2^k} + \frac{\tilde{r}}{2^k n} - x)^N}{N!}\right) \\
&+ \frac{(2a+1)^N}{2^{kN} N!}\left(D(f^{(N)}(x), \tilde{o}) + \omega_1^{(\mathcal{F})}\left(f^{(N)}, \frac{a}{2^k}\right)\right). \qquad (18)
\end{aligned}$$

Next we observe, as in the proof of Theorem 2, that

$$\begin{aligned}
&D\left(\mathcal{R}_N\left(x - \frac{a}{2^k}, \frac{j}{2^k} + \frac{\tilde{r}}{2^k n}\right), f^{(N)}\left(x - \frac{a}{2^k}\right) \odot \frac{(\frac{j+a}{2^k} + \frac{\tilde{r}}{2^k n} - x)^N}{N!}\right) \\
&\leq \frac{(2a+1)^N}{2^{kN} N!}\omega_1^{(\mathcal{F})}\left(f^{(N)}, \frac{2a+1}{2^k}\right). \qquad (19)
\end{aligned}$$

Hence the previous inequality (19) implies

$$\begin{aligned}
&D\left(\mathcal{R}_N\left(x - \frac{a}{2^k}, \frac{j}{2^k} + \frac{\tilde{r}}{2^k n}\right), \tilde{o}\right) \\
&\leq \frac{(2a+1)^N}{2^{kN} N!}\left(D(f^{(N)}(x), \tilde{o}) + 3\omega_1^{(\mathcal{F})}\left(f^{(N)}, \frac{a}{2^k}\right) + \omega_1^{(\mathcal{F})}\left(f^{(N)}, \frac{1}{2^k}\right)\right). \qquad (20)
\end{aligned}$$

Using the last inequality (20) into $(*)$ we obtain

$$\begin{aligned}
(*) &\leq \sum_{i=1}^{N-1} \frac{(2a+1)^i}{2^{ki} i!}\left(D(f^{(i)}(x), \tilde{o}) + \omega_1^{(\mathcal{F})}\left(f^{(i)}, \frac{a}{2^k}\right)\right) \\
&+ \frac{(2a+1)^N}{2^{kN} N!}\left(D(f^{(N)}(x), \tilde{o}) + 3\omega_1^{(\mathcal{F})}\left(f^{(N)}, \frac{a}{2^k}\right) + \omega_1^{(\mathcal{F})}\left(f^{(N)}, \frac{1}{2^k}\right)\right). \qquad (21)
\end{aligned}$$

By (21) the proof of the theorem is now finished. □

We need also the following main result.

Theorem 4. *All assumptions are as in Theorem 2. Define for $k \in \mathbb{Z}$, $x \in \mathbb{R}$ the fuzzy wavelet type operator*

$$(C_k f)(x) := \sum_{j=-\infty}^{\infty}{}^{*} \left(2^k \odot (FR)\int_0^{2^{-k}} f\left(t + \frac{j}{2^k}\right) dt\right) \odot \varphi(2^k x - j). \qquad (22)$$

Then

$$
\begin{aligned}
&D\left((C_k f)(x), f\left(x-\frac{a}{2^k}\right)\right)\\
&\quad\leq \sum_{i=1}^{N-1}\frac{(2a+1)^i}{2^{ki}i!}\left(D(f^{(i)}(x),\tilde{o})+\omega_1^{(\mathcal{F})}\left(f^{(i)},\frac{a}{2^k}\right)\right)\\
&\quad+\frac{(2a+1)^N}{2^{kN}N!}\left(D(f^{(N)}(x),\tilde{o})+3\omega_1^{(\mathcal{F})}\left(f^{(N)},\frac{a}{2^k}\right)+\omega_1^{(\mathcal{F})}\left(f^{(N)},\frac{1}{2^k}\right)\right)=\Delta_k\left(\frac{a}{2^k}\right),
\end{aligned}
\tag{23}
$$

all $k\in\mathbb{Z}$, $x\in\mathbb{R}$.

If $f\in C_{\mathcal{F}}^{NU}(\mathbb{R})$ *and as* $k\to+\infty$ *we obtain with rates that*

$$\lim_{k\to+\infty} D\left((C_k f)(x), f\left(x-\frac{a}{2^k}\right)\right)=0.$$

Corollary 5 (to Theorem 4, for $N=1$). *It holds*

$$
\begin{aligned}
D\left((C_k f)(x), f\left(x-\frac{a}{2^k}\right)\right)\leq\left(\frac{a}{2^{k-1}}+\frac{1}{2^k}\right)&(D(f'(x),\tilde{o})\\
&+3\omega_1^{(\mathcal{F})}\left(f',\frac{a}{2^k}\right)+\omega_1^{(\mathcal{F})}\left(f',\frac{1}{2^k}\right)\right),
\end{aligned}
$$

for any $x\in\mathbb{R}$, $k\in\mathbb{Z}$.

Corollary 6 (to Theorem 4). *The following improvement of (23) holds*

$$D\left((C_k f)(x), f\left(x-\frac{a}{2^k}\right)\right)\leq\min\left(\left(\omega_1^{(\mathcal{F})}\left(f,\frac{a+1}{2^k}\right)+\omega_1^{(\mathcal{F})}\left(f,\frac{a}{2^k}\right)\right),\Delta_k\left(\frac{a}{2^k}\right)\right),$$

for any $x\in\mathbb{R}$, $k\in\mathbb{Z}$.

Proof. From Theorem 3, inequality (10), of [4] we get

$$D((C_k f)(x), f(x))\leq\omega_1^{(\mathcal{F})}\left(f,\frac{a+1}{2^k}\right).$$

Then we observe that

$$
\begin{aligned}
D\left((C_k f)(x), f\left(x-\frac{a}{2^k}\right)\right)&\leq D((C_k f)(x), f(x))+D\left(f(x), f\left(x-\frac{a}{2^k}\right)\right)\\
&\leq\omega_1^{(\mathcal{F})}\left(f,\frac{a+1}{2^k}\right)+\omega_1^{(\mathcal{F})}\left(f,\frac{a}{2^k}\right). \qquad\square
\end{aligned}
$$

Proof of Theorem 4. Because φ is of compact support in $[-a,a]$ we see that

$$(C_k f)(x)=\sum_{\substack{j\\ 2^kx-j\in[-a,a]}}^{*}\left(2^k\odot(FR)\int_0^{2^{-k}} f\left(t+\frac{j}{2^k}\right)dt\right)\odot\varphi(2^kx-j). \tag{24}$$

Hence for specific $x \in \mathbb{R}$, $k \in \mathbb{Z}$ we have the j's in (24) to satisfy

$$x - \frac{a}{2^k} \leq \frac{j}{2^k} \leq x + \frac{a}{2^k}.$$

Using the fuzzy Taylor formula (Theorem 1, [3]) we find

$$f\left(t + \frac{j}{2^k}\right) = \sum_{i=0}^{N-1}{}^{*} f^{(i)}\left(x - \frac{a}{2^k}\right) \odot \frac{(t + \frac{(j+a)}{2^k} - x)^i}{i!} \oplus \mathcal{R}_N\left(x - \frac{a}{2^k}, t + \frac{j}{2^k}\right), \tag{25}$$

where

$$\mathcal{R}_N\left(x - \frac{a}{2^k}, t + \frac{j}{2^k}\right) := (FR)\int_{x-\frac{a}{2^k}}^{t+\frac{j}{2^k}}\left(\int_{x-\frac{a}{2^k}}^{s_1} \cdots\left(\left(\int_{x-\frac{a}{2^k}}^{s_{N-1}} f^{(N)}(s_N)ds_N\right)ds_{N-1}\cdots\right)ds_1. \tag{26}$$

Then by Theorem 3.4, [8] we get

$$2^k \odot (FR)\int_0^{2^{-k}} f\left(t + \frac{j}{2^k}\right)dt = \sum_{i=0}^{N-1}{}^{*} f^{(i)}\left(x - \frac{a}{2^k}\right) \odot \frac{2^k}{i!}\int_0^{2^{-k}}\left(t + \frac{(j+a)}{2^k} - x\right)^i dt$$
$$\oplus\, 2^k \odot (FR)\int_0^{2^{-k}} \mathcal{R}_N\left(x - \frac{a}{2^k}, t + \frac{j}{2^k}\right)dt. \tag{27}$$

Thus we have

$$(C_k f)(x) = \sum_{i=0}^{N-1}{}^{*} f^{(i)}\left(x - \frac{a}{2^k}\right) \odot \sum_{\substack{j \\ 2^k x - j \in [-a,a]}} \varphi(2^k x - j)\left(\frac{2^k}{i!}\int_0^{2^{-k}}\left(t + \frac{(j+a)}{2^k} - x\right)^i dt\right)$$
$$\oplus \sum_{\substack{j \\ 2^k x - j \in [-a,a]}} \varphi(2^k x - j)\left(2^k \odot (FR)\int_0^{2^{-k}} \mathcal{R}_N\left(x - \frac{a}{2^k}, t + \frac{j}{2^k}\right)dt\right). \tag{28}$$

Next we estimate

$$D\left((C_k f)(x), f\left(x - \frac{a}{2^k}\right)\right)$$
$$= D\left((C_k f)(x), \sum_{\substack{j \\ 2^k x - j \in [-a,a]}}^{*} \varphi(2^k x - j) \odot f\left(x - \frac{a}{2^k}\right)\right)$$
$$= D\left(\left(\sum_{i=1}^{N-1}{}^{*} f^{(i)}\left(x - \frac{a}{2^k}\right) \odot \sum_{\substack{j \\ 2^k x - j \in [-a,a]}} \varphi(2^k x - j)\left(\frac{2^k}{i!}\int_0^{2^{-k}}\left(t + \frac{(j+a)}{2^k} - x\right)^i dt\right)\right.\right.$$

$$\oplus \sum_{\substack{j \\ 2^k x - j \in [-a,a]}} \varphi(2^k x - j) \left(2^k \odot (FR) \int_0^{2^{-k}} \mathcal{R}_N \left(x - \frac{a}{2^k}, t + \frac{j}{2^k} \right) dt \right) \Bigg) , \tilde{o} \Bigg)$$

$$\leq \sum_{i=1}^{N-1} \sum_{\substack{j \\ 2^k x - j \in [-a,a]}} \varphi(2^k x - j) \left(\frac{2^k}{i!} \int_0^{2^{-k}} \left(t + \frac{(j+a)}{2^k} - x \right)^i dt \right) D \left(f^{(i)} \left(x - \frac{a}{2^k} \right), \tilde{o} \right)$$

$$+ \sum_{\substack{j \\ 2^k x - j \in [-a,a]}} \varphi(2^k x - j) \left(2^k D \left((FR) \int_0^{2^{-k}} \mathcal{R}_N \left(x - \frac{a}{2^k}, t + \frac{j}{2^k} \right) dt, \tilde{o} \right) \right)$$

$$\overset{\text{(by Lemma 1, [3])}}{\leq} \sum_{i=1}^{N-1} \frac{(2a+1)^i}{2^{ki} i!} \left(D(f^{(i)}(x), \tilde{o}) + \omega_1^{(\mathcal{F})} \left(f^{(i)}, \frac{a}{2^k} \right) \right)$$

$$+ \sum_{\substack{j \\ 2^k x - j \in [-a,a]}} \varphi(2^k x - j) \left(2^k \int_0^{2^{-k}} D \left(\mathcal{R}_N \left(x - \frac{a}{2^k}, t + \frac{j}{2^k} \right), \tilde{o} \right) dt \right) =: (*).$$

Next we work on

$$\begin{aligned} &D \left(\mathcal{R}_N \left(x - \frac{a}{2^k}, t + \frac{j}{2^k} \right), \tilde{o} \right) \\ &\quad \leq D \left(\mathcal{R}_N \left(x - \frac{a}{2^k}, t + \frac{j}{2^k} \right), f^{(N)} \left(x - \frac{a}{2^k} \right) \odot \frac{(t + \frac{(j+a)}{2^k} - x)^N}{N!} \right) \\ &\qquad + D \left(f^{(N)} \left(x - \frac{a}{2^k} \right) \odot \frac{(t + \frac{(j+a)}{2^k} - x)^N}{N!}, \tilde{o} \right). \end{aligned}$$

So that

$$\begin{aligned} &D \left(\mathcal{R}_N \left(x - \frac{a}{2^k}, t + \frac{j}{2^k} \right), \tilde{o} \right) \\ &\quad \leq D \left(\mathcal{R}_N \left(x - \frac{a}{2^k}, t + \frac{j}{2^k} \right), f^{(N)} \left(x - \frac{a}{2^k} \right) \odot \frac{(t + \frac{(j+a)}{2^k} - x)^N}{N!} \right) \\ &\qquad + \frac{(2a+1)^N}{2^{kN} N!} \left(D(f^{(N)}(x), \tilde{o}) + \omega_1^{(\mathcal{F})} \left(f^{(N)}, \frac{a}{2^k} \right) \right), \quad \text{for } 0 \leq t \leq 2^{-k}. \end{aligned} \tag{29}$$

Next, we act as in the proof of Theorem 2, we observe that

$$\begin{aligned} &D \left(\mathcal{R}_N \left(x - \frac{a}{2^k}, t + \frac{j}{2^k} \right), f^{(N)} \left(x - \frac{a}{2^k} \right) \odot \frac{(t + \frac{(j+a)}{2^k} - x)^N}{N!} \right) \\ &\qquad \leq \frac{(2a+1)^N}{2^{kN} N!} \omega_1^{(\mathcal{F})} \left(f^{(N)}, \frac{2a+1}{2^k} \right), \quad \text{for } 0 \leq t \leq 2^{-k}. \end{aligned} \tag{30}$$

Therefore we obtain

$$D\left(\mathcal{R}_N\left(x-\frac{a}{2^k},t+\frac{j}{2^k}\right),\tilde{o}\right) \tag{31}$$
$$\leq \frac{(2a+1)^N}{2^{kN}N!}\left(D(f^{(N)}(x),\tilde{o})+3\omega_1^{(\mathcal{F})}\left(f^{(N)},\frac{a}{2^k}\right)+\omega_1^{(\mathcal{F})}\left(f^{(N)},\frac{1}{2^k}\right)\right),$$

for $0\leq t\leq 2^{-k}$. Using the last inequality (31) into $(*)$ we get

$$(*)\leq\sum_{i=1}^{N-1}\frac{(2a+1)^i}{2^{ki}i!}\left(D(f^{(i)}(x),\tilde{o})+\omega_1^{(\mathcal{F})}\left(f^{(i)},\frac{a}{2^k}\right)\right)$$
$$+\frac{(2a+1)^N}{2^{kN}N!}\left(D(f^{(N)}(x),\tilde{o})+3\omega_1^{(\mathcal{F})}\left(f^{(N)},\frac{a}{2^k}\right)+\omega_1^{(\mathcal{F})}\left(f^{(N)},\frac{1}{2^k}\right)\right). \tag{32}$$

By (32) we have the validity of the theorem. □

The previous results lead to the following important theorems.

Theorem 5. *All assumptions are as in Theorem 2. It holds*

$$D((B_kf)(x),(D_kf)(x))\leq\beta_k\left(\frac{a}{2^k}\right)+\Delta_k\left(\frac{a}{2^k}\right), \tag{33}$$

for any $x\in\mathbb{R}$, $k\in\mathbb{Z}$. If $f\in C_{\mathcal{F}}^{NU}(\mathbb{R})$, as $k\to+\infty$, we obtain with rates that

$$\lim_{k\to+\infty}D((B_kf)(x),(D_kf)(x))=0.$$

Proof. By

$$D((B_kf)(x),(D_kf)(x))\leq D\left((B_kf)(x),f\left(x-\frac{a}{2^k}\right)\right)+D\left((D_kf)(x),f\left(x-\frac{a}{2^k}\right)\right)$$
$$\overset{\text{(by (2) and (12))}}{\leq}\beta_k\left(\frac{a}{2^k}\right)+\Delta_k\left(\frac{a}{2^k}\right).$$ □

Theorem 6. *All assumptions are as in Theorem 2. It holds*

$$D((D_kf)(x),(C_kf)(x))\leq 2\Delta_k\left(\frac{a}{2^k}\right), \tag{34}$$

for any $x\in\mathbb{R}$, $k\in\mathbb{Z}$. If $f\in C_{\mathcal{F}}^{NU}(\mathbb{R})$, as $k\to+\infty$, we get with rates that

$$\lim_{k\to+\infty}D((D_kf)(x),(C_kf)(x))=0.$$

Proof. By

$$D((D_kf)(x),(C_kf)(x))\leq D\left((D_kf)(x),f\left(x-\frac{a}{2^k}\right)\right)+D\left((C_kf)(x),f\left(x-\frac{a}{2^k}\right)\right)$$
$$\overset{\text{(by (12) and (23))}}{\leq}\Delta_k\left(\frac{a}{2^k}\right)+\Delta_k\left(\frac{a}{2^k}\right)=2\Delta_k\left(\frac{a}{2^k}\right).$$ □

Theorem 7. *All assumptions are as in Theorem 2. It holds*

$$D((B_k f)(x),(C_k f)(x)) \leq \beta_k\left(\frac{a}{2^k}\right)+\Delta_k\left(\frac{a}{2^k}\right), \tag{35}$$

all $x\in\mathbb{R}$, $k\in\mathbb{Z}$. If $f\in C_{\mathcal{F}}^{NU}(\mathbb{R})$, as $k\to+\infty$, we find with rates that

$$\lim_{k\to+\infty} D((B_k f)(x),(C_k f)(x))=0.$$

Proof. By

$$D((B_k f)(x),(C_k f)(x)) \leq D\left((B_k f)(x), f\left(x-\frac{a}{2^k}\right)\right)+D\left((C_k f)(x), f\left(x-\frac{a}{2^k}\right)\right) \overset{\text{(by (2) and (23))}}{\leq} \beta_k\left(\frac{a}{2^k}\right)+\Delta_k\left(\frac{a}{2^k}\right). \qquad \square$$

It follows another family of basic interesting related results.

Theorem 8. *Here φ is as in Theorem 2 and $f\in C(\mathbb{R},\mathbb{R}_{\mathcal{F}})$. Fuzzy wavelet type operators B_k defined by (1), and C_k defined by (22). Then*

$$D((B_k f)(x),(C_k f)(x)) \leq \omega_1^{(\mathcal{F})}\left(f,\frac{1}{2^k}\right), \tag{36}$$

for any $x\in\mathbb{R}$, $k\in\mathbb{Z}$. If $f\in C_{\mathcal{F}}^{U}(\mathbb{R})$, as $k\to+\infty$, we obtain with rates that

$$\lim_{k\to+\infty} D((B_k f)(x),(C_k f)(x))=0.$$

Proof. We have that

$$\begin{aligned}
&D((B_k f)(x),(C_k f)(x)) \\
&= D\left(\sum_{j=-\infty}^{\infty}{}^{*} f\left(\frac{j}{2^k}\right)\odot\varphi(2^k x-j), \sum_{j=-\infty}^{\infty}{}^{*}\left(2^k\odot(FR)\int_0^{2^{-k}} f\left(t+\frac{j}{2^k}\right)dt\right)\odot\varphi(2^k x-j)\right) \\
&\leq \sum_{\substack{j\\ 2^k x-j\in[-a,a]}} \varphi(2^k x-j) D\left(f\left(\frac{j}{2^k}\right), 2^k\odot(FR)\int_0^{2^{-k}} f\left(t+\frac{j}{2^k}\right)dt\right) \\
&= 2^k \sum_{\substack{j\\ 2^k x-j\in[-a,a]}} \varphi(2^k x-j) D\left((FR)\int_0^{2^{-k}} f\left(\frac{j}{2^k}\right)dt, (FR)\int_0^{2^{-k}} f\left(t+\frac{j}{2^k}\right)dt\right) \\
&\overset{\text{(by Lemma 1, [3])}}{\leq} 2^k \sum_{\substack{j\\ 2^k x-j\in[-a,a]}} \varphi(2^k x-j)\int_0^{2^{-k}} D\left(f\left(\frac{j}{2^k}\right), f\left(t+\frac{j}{2^k}\right)\right)dt \\
&\leq 2^k \sum_{\substack{j\\ 2^k x-j\in[-a,a]}} \varphi(2^k x-j)\int_0^{2^{-k}} \omega_1^{(\mathcal{F})}(f,t)dt \leq \omega_1^{(\mathcal{F})}\left(f,\frac{1}{2^k}\right). \qquad \square
\end{aligned}$$

Theorem 9. *Here φ is as in Theorem 2 and $f \in C(\mathbb{R}, \mathbb{R}_{\mathcal{F}})$. Fuzzy wavelet type operators D_k defined by (10) and (11), and C_k defined by (22). Then*

$$D((D_k f)(x), (C_k f)(x)) \leq \omega_1^{(\mathcal{F})}\left(f, \frac{1}{2^{k-1}}\right), \tag{37}$$

for any $x \in \mathbb{R}$, $k \in \mathbb{Z}$. If $f \in C_{\mathcal{F}}^U(\mathbb{R})$, as $k \to +\infty$, we get with rates that

$$\lim_{k \to +\infty} D((D_k f)(x), (C_k f)(x)) = 0.$$

Proof. We have that

$$\begin{aligned}
&D((D_k f)(x), (C_k f)(x)) \\
&= D\left(\sum_{j=-\infty}^{\infty}{}^{*} \delta_{kj}(f) \odot \varphi(2^k x - j), \sum_{j=-\infty}^{\infty}{}^{*} \left(2^k \odot (FR)\int_0^{2^{-k}} f\left(t + \frac{j}{2^k}\right) dt\right) \odot \varphi(2^k x - j)\right) \\
&\leq \sum_{\substack{j \\ 2^k x - j \in [-a,a]}} \varphi(2^k x - j) D\left(\delta_{kj}(f), 2^k \odot (FR)\int_0^{2^{-k}} f\left(t + \frac{j}{2^k}\right) dt\right) \\
&\leq \sum_{\substack{j \\ 2^k x - j \in [-a,a]}} \varphi(2^k x - j) \sum_{\tilde{r}=0}^{n} w_{\tilde{r}} D\left(f\left(\frac{j}{2^k} + \frac{\tilde{r}}{2^k n}\right), 2^k \odot (FR)\int_0^{2^{-k}} f\left(t + \frac{j}{2^k}\right) dt\right) \\
&= 2^k \sum_{\substack{j \\ 2^k x - j \in [-a,a]}} \varphi(2^k x - j) \sum_{\tilde{r}=0}^{n} w_{\tilde{r}} D\left((FR)\int_0^{2^{-k}} f\left(\frac{j}{2^k} + \frac{\tilde{r}}{2^k n}\right) dt,\right. \\
&\qquad\qquad \left.(FR)\int_0^{2^{-k}} f\left(t + \frac{j}{2^k}\right) dt\right) \\
&\overset{\text{(by Lemma 1, [3])}}{\leq} 2^k \sum_{\substack{j \\ 2^k x - j \in [-a,a]}} \varphi(2^k x - j) \sum_{\tilde{r}=0}^{n} w_{\tilde{r}} \int_0^{2^{-k}} D\left(f\left(\frac{j}{2^k} + \frac{\tilde{r}}{2^k n}\right), f\left(t + \frac{j}{2^k}\right)\right) dt \\
&\leq 2^k \sum_{\substack{j \\ 2^k x - j \in [-a,a]}} \varphi(2^k x - j) \sum_{\tilde{r}=0}^{n} w_{\tilde{r}} \int_0^{2^{-k}} \omega_1^{(\mathcal{F})}\left(f, \left|\frac{\tilde{r}}{2^k n} - t\right|\right) dt \\
&\leq 2^k \sum_{\substack{j \\ 2^k x - j \in [-a,a]}} \varphi(2^k x - j) \sum_{\tilde{r}=0}^{n} w_{\tilde{r}} \int_0^{2^{-k}} \omega_1^{(\mathcal{F})}\left(f, \frac{1}{2^{k-1}}\right) dt = \omega_1^{(\mathcal{F})}\left(f, \frac{1}{2^{k-1}}\right). \qquad \square
\end{aligned}$$

Theorem 10. *Here φ is as in Theorem 2 and $f \in C(\mathbb{R}, \mathbb{R}_{\mathcal{F}})$. Fuzzy wavelet type operators B_k defined by (1), and D_k defined by (10) and (11). Then*

$$D((B_k f)(x), (D_k f)(x)) \leq \omega_1^{(\mathcal{F})}\left(f, \frac{1}{2^k}\right), \tag{38}$$

for any $x \in \mathbb{R}$, $k \in \mathbb{Z}$. *If* $f \in C^U_{\mathcal{F}}(\mathbb{R})$, *as* $k \to +\infty$, *we find with rates that*

$$\lim_{k \to +\infty} D((B_k f)(x), (D_k f)(x)) = 0.$$

Proof. We see that

$$\begin{aligned}
&D((B_k f)(x), (D_k f)(x)) \\
&= D\left(\sum_{j=-\infty}^{\infty}{}^{*} f\left(\frac{j}{2^k}\right) \odot \varphi(2^k x - j), \sum_{j=-\infty}^{\infty}{}^{*} \delta_{kj}(f) \odot \varphi(2^k x - j) \right) \\
&\leq \sum_{\substack{j \\ 2^k x - j \in [-a,a]}} \varphi(2^k x - j) D\left(f\left(\frac{j}{2^k}\right), \delta_{kj}(f) \right) \\
&\leq \sum_{\substack{j \\ 2^k x - j \in [-a,a]}} \varphi(2^k x - j) \sum_{\tilde{r}=0}^{n} w_{\tilde{r}} D\left(f\left(\frac{j}{2^k}\right), f\left(\frac{j}{2^k} + \frac{\tilde{r}}{2^k n}\right) \right) \\
&\leq \sum_{\substack{j \\ 2^k x - j \in [-a,a]}} \varphi(2^k x - j) \sum_{\tilde{r}=0}^{n} w_{\tilde{r}} \omega_1^{(\mathcal{F})}\left(f, \frac{\tilde{r}}{2^k n}\right) \leq \omega_1^{(\mathcal{F})}\left(f, \frac{1}{2^k}\right). \qquad \square
\end{aligned}$$

Next we present the corresponding comparison results based on the previously given theorems.

Corollary 7 (to Theorem 5, and Theorem 10). *All assumptions here are as in Theorem 2. It holds*

$$D((B_k f)(x), (D_k f)(x)) \leq \min\left(\omega_1^{(\mathcal{F})}\left(f, \frac{1}{2^k}\right), \beta_k\left(\frac{a}{2^k}\right) + \Delta_k\left(\frac{a}{2^k}\right) \right), \tag{39}$$

for any $x \in \mathbb{R}$, $k \in \mathbb{Z}$.

Corollary 8 (to Theorem 6 and Theorem 9). *All assumptions here are as in Theorem 2. It holds*

$$D((D_k f)(x), (C_k f)(x)) \leq \min\left(\omega_1^{(\mathcal{F})}\left(f, \frac{1}{2^{k-1}}\right), 2\Delta_k\left(\frac{a}{2^k}\right) \right), \tag{40}$$

for any $x \in \mathbb{R}$, $k \in \mathbb{Z}$.

Corollary 9 (to Theorem 7 and Theorem 8). *All assumptions here are as in Theorem 2. It holds*

$$D((B_k f)(x), (C_k f)(x)) \leq \min\left(\omega_1^{(\mathcal{F})}\left(f, \frac{1}{2^k}\right), \beta_k\left(\frac{a}{2^k}\right) + \Delta_k\left(\frac{a}{2^k}\right) \right), \tag{41}$$

all $x \in \mathbb{R}$, $k \in \mathbb{Z}$.

Finally we study similarly the Fuzzy wavelet operator A_k.

Theorem 11. *Let* $f \in C_b(\mathbb{R}, \mathbb{R}_{\mathcal{F}})$ *and the scaling function* $\varphi(x)$ *a real valued function with* supp $\varphi(x) \subseteq [-a, a]$, $0 < a < +\infty$, φ *is continuous on* $[-a, a]$, $\varphi(x) \geq 0$, *such that* $\sum\limits_{j=-\infty}^{\infty} \varphi(x-1) = 1$ *on* $\mathbb{R}$ *(then* $\int_{-\infty}^{\infty} \varphi(x)dx = 1$*). Define*

$$\varphi_{kj}(t) := 2^{k/2}\varphi(2^k t - j), \quad \text{for } k, j \in \mathbb{Z},\ t \in \mathbb{R}, \tag{42}$$

$$\langle f, \varphi_{kj} \rangle := (FR)\int_{\frac{j-a}{2^k}}^{\frac{j+a}{2^k}} f(t) \odot \varphi_{kj}(t)dt, \tag{43}$$

and set

$$(A_k f)(x) := \sum_{j=-\infty}^{\infty}{}^{*} \langle f, \varphi_{kj} \rangle \odot \varphi_{kj}(x), \quad \text{any } x \in \mathbb{R}. \tag{44}$$

The fuzzy wavelet type operator $(B_k f)(x)$ *is defined by (1). Then it holds*

$$D((A_k f)(x), (B_k f)(x)) \leq \omega_1^{(\mathcal{F})}\left(f, \frac{a}{2^k}\right), \tag{45}$$

for any $x \in \mathbb{R}$, $k \in \mathbb{Z}$. *If* $f \in C_{\mathcal{F}}^U(\mathbb{R})$ *and bounded, then*

$$\lim_{k \to +\infty} D((A_k f)(x), (B_k f)(x)) = 0 \quad \text{with rates.}$$

Proof. We see easily that

$$(A_k f)(x) = \sum_{j=-\infty}^{\infty}{}^{*} \left((FR)\int_{j-a}^{j+a} f\left(\frac{u}{2^k}\right) \odot \varphi(u-j)du\right) \odot \varphi(2^k x - j). \tag{46}$$

Also it holds

$$\int_{j-a}^{j+a} \varphi(u-j)du = 1. \tag{47}$$

So we observe

$$\begin{aligned}
&D((A_k f)(x), (B_k f)(x)) \\
&= D\left(\sum_{j=-\infty}^{\infty}{}^{*} \left((FR)\int_{j-a}^{j+a} f\left(\frac{u}{2^k}\right) \odot \varphi(u-j)du\right)\right. \\
&\quad \left. \odot \varphi(2^k x - j), \sum_{j=-\infty}^{\infty}{}^{*} f\left(\frac{j}{2^k}\right) \odot \varphi(2^k x - j)\right) \\
&\leq \sum_{\substack{j \\ 2^k x - j \in [-a,a]}} \varphi(2^k x - j) D\left((FR)\int_{j-a}^{j+a} f\left(\frac{u}{2^k}\right) \odot \varphi(u-j)du,\right.
\end{aligned}$$

$$(FR)\int_{j-a}^{j+a} f\left(\frac{j}{2^k}\right)\odot\varphi(u-j)du\bigg)$$

$$\overset{\substack{\text{(by Lemma 1, [3]}\\ \text{and Lemma 2, [4])}}}{\leq} \sum_{\substack{j\\ 2^kx-j\in[-a,a]}} \varphi(2^kx-j)\int_{j-a}^{j+a}\varphi(u-j)D\left(f\left(\frac{u}{2^k}\right), f\left(\frac{j}{2^k}\right)\right)du$$

$$\leq \sum_{\substack{j\\ 2^kx-j\in[-a,a]}} \varphi(2^kx-j)\int_{j-a}^{j+a}\varphi(u-j)\omega_1^{(\mathcal{F})}\left(f,\frac{a}{2^k}\right)du = \omega_1^{(\mathcal{F})}\left(f,\frac{a}{2^k}\right). \qquad \square$$

Theorem 12. *Let φ, f, A_k as in Theorem 11. Let fuzzy wavelet type operator D_k as in (10) and (11). Then*

$$D((A_kf)(x),(D_kf)(x)) \leq \omega_1^{(\mathcal{F})}\left(f,\frac{a+1}{2^k}\right), \tag{48}$$

for any $x\in\mathbb{R}$, $k\in\mathbb{Z}$. If $f\in C_{\mathcal{F}}^U(\mathbb{R})$ and bounded then

$$\lim_{k\to+\infty} D((A_kf)(x),(D_kf)(x)) = 0 \quad \textit{with rates.}$$

Proof. We notice that

$$D((A_kf)(x),(D_kf)(x))$$

$$\overset{\text{(by (46))}}{=} D\left(\sum_{j=-\infty}^{\infty}{}^{*}\left((FR)\int_{j-a}^{j+a} f\left(\frac{u}{2^k}\right)\odot\varphi(u-j)du\right)\right.$$

$$\left.\cdot\,\varphi(2^kx-j), \sum_{j=-\infty}^{\infty}{}^{*}\delta_{kj}(f)\odot\varphi(2^kx-j)\right)$$

$$\leq \sum_{\substack{j\\ 2^kx-j\in[-a,a]}} \varphi(2^kx-j)D\left((FR)\int_{j-a}^{j+a} f\left(\frac{u}{2^k}\right)\right.$$

$$\left.\odot\,\varphi(u-j)du, \sum_{\tilde{r}=0}^{n}{}^{*} w_{\tilde{r}}\odot f\left(\frac{j}{2^k}+\frac{\tilde{r}}{2^kn}\right)\right)$$

$$\overset{\text{(by (47))}}{\leq} \sum_{\substack{j\\ 2^kx-j\in[-a,a]}} \varphi(2^kx-j)\sum_{\tilde{r}=0}^{n} w_{\tilde{r}}D\left((FR)\int_{j-a}^{j+a} f\left(\frac{u}{2^k}\right)\odot\varphi(u-j)du,\right.$$

$$\left.(FR)\int_{j-a}^{j+a} f\left(\frac{j}{2^k}+\frac{\tilde{r}}{2^kn}\right)\odot\varphi(u-j)du\right)$$

$$\overset{\substack{\text{(by Lemma 1, [3]}\\ \text{and Lemma 2, [4])}}}{\leq} \sum_{\substack{j\\ 2^kx-j\in[-a,a]}} \varphi(2^kx-j)\sum_{\tilde{r}=0}^{n} w_{\tilde{r}}$$

$$\times\left(\int_{j-a}^{j+a}\varphi(u-j)D\left(f\left(\frac{u}{2^k}\right),f\left(\frac{j}{2^k}+\frac{\tilde{r}}{2^k n}\right)\right)du\right)$$
$$\leq \sum_{\substack{j\\2^k x-j\in[-a,a]}}\varphi(2^k x-j)\sum_{\tilde{r}=0}^{n}w_{\tilde{r}}\left(\int_{j-a}^{j+a}\varphi(u-j)\omega_1^{(\mathcal{F})}\left(f,\frac{a+1}{2^k}\right)du\right)$$
$$=\omega_1^{(\mathcal{F})}\left(f,\frac{a+1}{2^k}\right). \qquad \square$$

Theorem 13. *Let φ, f, A_k as in Theorem 11. Let fuzzy wavelet type operator C_k as in (22). Then*

$$D((A_k f)(x),(C_k f)(x))\leq\omega_1^{(\mathcal{F})}\left(f,\frac{a+1}{2^k}\right), \tag{49}$$

for any $x\in\mathbb{R}$, $k\in\mathbb{Z}$. If $f\in C_{\mathcal{F}}^U(\mathbb{R})$ and bounded, then

$$\lim_{k\to+\infty}D((A_k f)(x),(C_k f)(x))=0 \quad \text{with rates.}$$

Proof. We observe that

$$D((A_k f)(x),(C_k f)(x))$$
$$\overset{\text{(by (46) and (22))}}{=} D\left(\sum_{j=-\infty}^{\infty}{}^{*}\left((FR)\int_{j-a}^{j+a}f\left(\frac{u}{2^k}\right)\odot\varphi(u-j)du\right)\odot\varphi(2^k x-j),\right.$$
$$\left.\sum_{j=-\infty}^{\infty}{}^{*}\left(2^k\odot(FR)\int_0^{2^{-k}}f\left(t+\frac{j}{2^k}\right)dt\right)\odot\varphi(2^k x-j)\right)$$
$$\leq \sum_{\substack{j\\2^k x-j\in[-a,a]}}\varphi(2^k x-j)D\left((FR)\int_{j-a}^{j+a}f\left(\frac{u}{2^k}\right)\right.$$
$$\left.\odot\,\varphi(u-j)du,2^k\odot(FR)\int_0^{2^{-k}}f\left(t+\frac{j}{2^k}\right)dt\right)$$
$$=\sum_{\substack{j\\2^k x-j\in[-a,a]}}\varphi(2^k x-j)D\left((FR)\int_{j-a}^{j+a}f\left(\frac{u}{2^k}\right)\odot\varphi(u-j)du,\right.$$
$$\left.(FR)\int_{j-a}^{j+a}\varphi(u-j)\odot\left[2^k\odot(FR)\int_0^{2^{-k}}f\left(t+\frac{j}{2^k}\right)dt\right]du\right)$$
$$\overset{\text{(by Lemma 1, [3] and Lemma 2, [4])}}{\leq}\sum_{\substack{j\\2^k x-j\in[-a,a]}}\varphi(2^k x-j)$$
$$\times\left(\int_{j-a}^{j+a}\varphi(u-j)D\left(f\left(\frac{u}{2^k}\right),2^k\odot(FR)\int_0^{2^{-k}}f\left(t+\frac{j}{2^k}\right)dt\right)du\right)$$

$$= \sum_{\substack{j \\ 2^k x - j \in [-a,a]}} \varphi(2^k x - j) \left(\int_{j-a}^{j+a} \varphi(u-j) 2^k D \right.$$
$$\left. \left((FR) \int_0^{2^{-k}} f\left(\frac{u}{2^k}\right) dt, (FR) \int_0^{2^{-k}} f\left(t + \frac{j}{2^k}\right) dt \right) du \right)$$

(by Lemma 1, [3])

$$\leq \sum_{\substack{j \\ 2^k x - j \in [-a,a]}} \varphi(2^k x - j) \left(\int_{j-a}^{j+a} \varphi(u-j) 2^k \right.$$
$$\left. \times \left(\int_0^{2^{-k}} D\left(f\left(\frac{u}{2^k}\right), f\left(t + \frac{j}{2^k}\right)\right) dt \right) du \right)$$
$$\leq \sum_{\substack{j \\ 2^k x - j \in [-a,a]}} \varphi(2^k x - j) \left(\int_{j-a}^{j+a} \varphi(u-j) 2^k \left(\int_0^{2^{-k}} \omega_1^{(\mathcal{F})} \left(f, \left| \frac{u}{2^k} - \frac{j}{2^k} - t \right| \right) dt \right) du \right)$$
$$\leq \sum_{\substack{j \\ 2^k x - j \in [-a,a]}} \varphi(2^k x - j) \left(\int_{j-a}^{j+a} \varphi(u-j) 2^k \left(\int_0^{2^{-k}} \omega_1^{(\mathcal{F})} \left(f, \frac{a+1}{2^k} \right) dt \right) du \right)$$
$$= \omega_1^{(\mathcal{F})} \left(f, \frac{a+1}{2^k} \right). \qquad \square$$

Example 1. The following scaling function φ fulfills the assumptions of the presented theorems

$$\varphi(x) = \begin{cases} x+1, & -1 \leq x \leq 0, \\ 1-x, & 0 < x \leq 1, \\ 0, & \text{elsewhere.} \end{cases}$$

Note 2. A_k, B_k, C_k, D_k are linear operators over $\mathbb{R}$.

Remark 1. On Theorems 8, 9, 10, 11, 12 and 13. It is enough to comment Theorem 8, similar conclusions can be derived from the rest of them. Assume that $f \in C(\mathbb{R}, \mathbb{R}_{\mathcal{F}})$ fulfills the following Lipschitz condition

$$D(f(x), f(y)) \leq M|x-y|^\rho, \quad 0 < \rho \leq 1,\ M > 0. \tag{50}$$

Then clearly it holds

$$\omega_1^{(\mathcal{F})} \left(f, \frac{1}{2^k} \right) \leq \frac{M}{2^{k\rho}}, \quad k \in \mathbb{Z}. \tag{51}$$

So that from (36) we obtain

$$D((B_k f)(x), (C_k f)(x)) \leq \frac{M}{2^{k\rho}}, \tag{52}$$

and $k \in \mathbb{Z}$, $x \in \mathbb{R}$. And consequently

$$D^*((B_k f),(C_k f)) \leq \frac{M}{2^{k\rho}}\,, \tag{53}$$

for any $k \in \mathbb{Z}$.

Finally we get the global error estimate

$$\sup_{\substack{f \in C(\mathbb{R},\mathbb{R}_{\mathcal{F}}):\\ f \text{ as in } (50)}} D^*((B_k f),(C_k f)) \leq \frac{M}{2^{k\rho}}\,. \tag{54}$$

Etc.

At the end we give two independent but related and useful results.

Proposition 1. *Let φ, f be both even functions as in Theorem 8. Let $(B_k f)$ defined by (1). Then $(B_k f)(x)$ is an even function.*

Proof. We observe for $x \in \mathbb{R}$ that

$$\begin{aligned}
(B_k f)(-x) &= \sum_{j=-\infty}^{\infty}{}^{*} f\left(\frac{j}{2^k}\right) \odot \varphi(-2^k x - j)\\
&= \sum_{j=-\infty}^{\infty}{}^{*} f\left(-\frac{j}{2^k}\right) \odot \varphi(2^k x + j) \quad \text{(a finite sum)}\\
&= \sum_{j^*:=-j=\infty}^{-\infty}{}^{*} f\left(\frac{j^*}{2^k}\right) \odot \varphi(2^k x - j^*)\\
&= \sum_{j^*=-\infty}^{\infty}{}^{*} f\left(\frac{j^*}{2^k}\right) \odot \varphi(2^k x - j^*) = (B_k f)(x). \qquad \square
\end{aligned}$$

Proposition 2. *Let φ, f be both even functions as in Theorem 11. Let $(A_k f)$ defined by (43) and (44). Then $(A_k f)(x)$ is an even function.*

Proof. We observe for $x \in \mathbb{R}$ that

$$\begin{aligned}
(A_k f)(-x) &\overset{(46)}{=} \sum_{j=-\infty}^{\infty}{}^{*} \left((FR)\int_{j-a}^{j+a} f\left(\frac{u}{2^k}\right) \odot \varphi(u-j)du\right) \odot \varphi(-2^k x - j)\\
&\qquad \text{(a finite sum)}\\
&= \sum_{j=-\infty}^{\infty}{}^{*} \left((FR)\int_{j-a}^{j+a} f\left(-\frac{u}{2^k}\right) \odot \varphi(-u+j)du\right) \odot \varphi(2^k x + j)\\
&\qquad \text{(linear change of variables is valid in } (FR)\text{-integrals)}\\
&= \sum_{j^*:=-j=\infty}^{-\infty}{}^{*} \left((FR)\int_{j^*-a}^{j^*+a} f\left(\frac{w}{2^k}\right) \odot \varphi(w-j^*)dw\right) \odot \varphi(2^k x - j^*)\\
&\overset{(46)}{=} (A_k f)(x). \qquad \square
\end{aligned}$$

References

[1] G. A. Anastassiou, Shape and probability preserving univariate wavelet type operators, *Commun. Appl. and Anal.*, **1**, No. 3 (1997), 303–314.

[2] G. A. Anastassiou, *Quantitative Approximations*, Chapman & Hall/CRC, Boca Raton, New York, 2001.

[3] G. A. Anastassiou, Rate of convergence of fuzzy neural network operators, univariate case, *Journal of Fuzzy Mathematics*, **10**, No. 3 (2002), 755–780.

[4] G. A. Anastassiou, Fuzzy wavelet type operators, submitted.

[5] G. A. Anastassiou and S. Gal, On a fuzzy trigonometric approximation theorem of Weierstrass-type, in *Journal of Fuzzy Mathematics*, **9**, No. 3 (2001), 701–708.

[6] G. A. Anastassiou and X.M. Yu, Monotone and probabilistic wavelet approximation, *Stochastic Anal. Appl.*, **10**(3) (1992), 251–264.

[7] S. Gal, Approximation theory in fuzzy setting. Chapter 13 in *Handbook of Analytic Computational Methods in Applied Mathematics* (edited by G. Anastassiou), Chapman & Hall CRC Press, Boca Raton, New York, 2000, pp. 617–666.

[8] R. Goetschel , Jr. and W. Voxman, Elementary fuzzy calculus, *Fuzzy Sets and Systems*, **18** (1986), 31–43.

[9] Congxin Wu and Zengtai Gong, On Henstock integral of fuzzy number valued functions (I), *Fuzzy Sets and Systems*, **120**, No. 3, 2001, 523–532.

[10] L. A. Zadeh, Fuzzy sets, *Information and Control*, **8**, 1965, 338–353.

JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS,VOL.5,NO.1,53-75,2007,COPYRIGHT 2007 EUDOXUS PRESS,LLC

Stochastic modeling for the COMET-assay

A.-L. Boulesteix* V. Hösel† V. Liebscher‡

April 19, 2006

Abstract

We present a stochastic model for single cell gel electrophoresis (COMET-assay) data. The distribution of length of DNA fragments is calculated according to a 'Random Breakage Model' and the migration of DNA fragments among gel fibers is discussed. Essential to our approach is the use of point process structures, renewal theory and reduction to intensity histograms for further data analysis. Parameter estimations and simulations illustrate the features of the model.

AMS Classification (2000):62P10 (Prim.); 92C40, 60K05, 60G55, 92C55

Keywords: COMET-assay, renewal process, point process, Poisson process, random breakage model

1 Introduction

Single cell gel electrophoresis or "COMET-assay" is a very efficient method to examine DNA damage and repair with many applications, for example in cancer research. A non-damaged DNA molecule is a long linear chain of desoxyribonucleic acids. When a cell is irradiated several strand breaks in the DNA may occur. The aim of the study is to

*Technical University of Munich, Department of Medical Statistics and Epidemiology, Technical University of Munich, Ismaningerstr.22, D–81675 Munich, Germany, anne-laure.boulesteix@tum.de

†Munich University of Technology, Centre of Mathematical Sciences, D–80539 Munich, Germany, hoesel@t-online.de

‡GSF — National Research Centre for Environment and Health, Institute of Biomathematics and Biometry, Ingolstädter Landstr.1, D–85758 Neuherberg, Germany, liebscher@gsf.de, *corresponding author*

detect to which amount a broken DNA molecule can be repaired by the organism. Non efficient repair may indicate genetically determined malfunctions in the recombination and replication mechanisms of the DNA. At present, the COMET assay is one of the most important techniques to monitor DNA damage and repair at the level of single cells.

The standard way to analyze COMET data is to compute characterizing geometric properties of the comet, e.g. the tail moment [3] or the comet moment [19]. Some of these parameters show only little variability across experiments [12]. However, all these parameters are sensible to small changes in the recorded COMET image. The used image processing method plays a major role and errors in detecting faster fragments may occur. Such fragments usually appear darker and are thus not well separable from the background of the image. Further, the images contain much more information than could be coded in one or a few parameters. A comprehensive model for the whole data and some robust methods to extract relevant information would allow to make better use of the recorded images. In the present work, we suggest such a modeling approach.

This work is structured as follows. Section 2 is a short introduction into the COMET-assay and our modeling approach. Section 3 describes the problem as a marked point process. In sections 4,5,6, we derive stochastic models for the various aspects of the experiment such as, e.g., the distribution of fragment masses after radiation or the mass dependent migration distance of a single DNA fragment. In section 7, we discuss the combined model by means of simulation and parameter estimation.

2 The COMET-assay and its modeling

A large amount of articles describe the technical details of the COMET-assay, for instance [3, 6, 19]. An up to date source of information is the web site [17] and an extensive review of the COMET-assay can be found in [13]. Here, we only present a short overview of the method with emphasis on a few features which are important for our mathematical modeling.

In COMET experiments, the cells to analyze are attached to an agarose gel and placed in an electric field, after suitable treatment and in particular conditions. Since DNA is polar, DNA molecules tend to migrate. Big DNA molecules (i.e. non-damaged or repaired DNA molecules) show no observable migration, whereas small DNA molecules (i.e. damaged DNA molecules) migrate quickly off the center of the cell. These small fragments

are responsible for the comet like shape of the electrophoresis image (tail), hence the name COMET-assay, see Figure 1. It is quite difficult to explain why small molecules migrate faster than big ones, but one of the main explanations is that big molecules are more sensitive to hurdles (gel fibers) during the migration. Till now there are diverse opinions among biologists about the underlying mechanisms. Anyway, at the end of the electrophoresis, it is possible to see whether a cell is 'quite damaged' or 'quite non-damaged', by analyzing the shape of the comet: a damaged cell has a long and/or dense tail, whereas a non-damaged cell merely looks like a homogeneous disk.

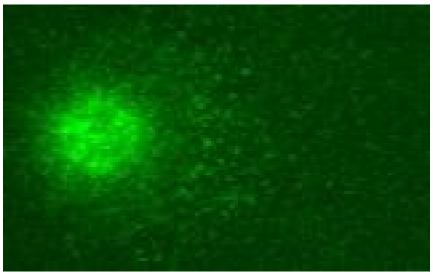


Figure 1: A comet from an irradiated cell

Our aim in the present paper is to establish a reasonable stochastic model describing the data which can serve as a basis for future statistical inference. This strategy is in contrast to the standard approach, which uses only a few geometric features. We emphasize again that interesting information contained in the image data is not encoded in the single geometric parameters. We are able to retrieve this information only if we can model the physical processes of the experiment with sufficient accuracy. The final goal is to estimate the distribution of lengths of DNA molecules (or equivalently their distribution of mass) in damaged and repaired cells, in order to get more information on the repair mechanism and its efficacy. Keeping track of the approximations and assumptions in the modeling process will help to implement methods which are robust under changes of model parameters and slight violations of model assumptions.

In terms of the data, the distribution of molecule lengths we want to estimate is best associated to the distribution of displacements of single DNA molecules. This demands some further knowledge about the relation between length and speed. In the literature, biologists propose theoretical models (for instance in [30]) for this relation and give experimental results (for instance in [23]) obtained in various conditions. These studies are especially designed for usual gel electrophoresis, where the lengths of the DNA fragments

is ≈ 500 bp. This is much smaller than the fragment lengths considered in COMET experiments. In section 3 we propose a global model to describe the DNA migration and finally get a formula agreeing with some of the experimental results found in the literature. Our model takes into account a great part of the physical features cited in the literature and is quite consistent with the empirical formulas already known.

Our model for the available data is guided by the experiment. First, we describe the placement of the DNA fragments before radiation and after radiation. To model the effect of the gel electrophoresis, we then give a mathematical description of the migration of DNA molecules through an agarose gel.

3 Marked Point Processes as Description

We consider a single cell containing N DNA fragments, where N is a (random) number depending on the number of DNA breaks. Each of the N fragments is represented by a tuple $(\mathbf{X}_i, m_i)$, $i \in \{1, \ldots, N\}$, where $\mathbf{X}_i$ is a three dimensional vector representing the initial location of fragment i and m_i is its mass. $\Xi = \{(\mathbf{X}_i, m_i) : i \in \{1, \ldots, N\}\}$ corresponds to the observed fragments, approximating the location of a fragment by a point, but carrying its mass into the calculations via m_i. Note, that we can not differentiate between break experiments resulting in fractions of the same size. So, the set Ξ which is a simple finite marked *point process* [11] is a natural description of the fragments. Let $\mathbf{D}_i$ be the three dimensional vector of displacement of the i-th fragment during the experiment and $\mathbf{X}'_i = \mathbf{X}_i + \mathbf{D}_i$, which is thus the three dimensional vector of end location of fragment i. $\mathbf{X}$, $\mathbf{D}$ and $\mathbf{X}'$ are depicted in figure 2 for one point.

4 The Length Distribution — Poisson Approximation

Our first goal is to determine the distribution of fraction lengths. It remains a very complex issue because we do not know much about the mechanisms of breakage and repair.

However, with a few simple assumptions, one can regard the distribution of lengths in a damaged cell as exponential. This model is commonly called 'Random Breakage Model' (RBM) and described in [21]. Let us briefly discuss its underlying assumptions.

1. Breaks occur completely at random, i.e. the radiation causes only single break events which do not influence each other (this hypothesis is supported by the low

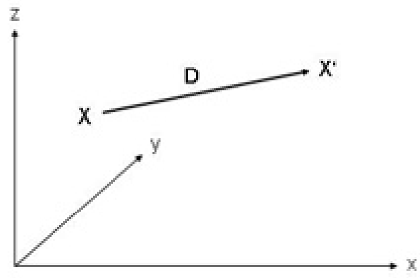


Figure 2: Coordinate system. The represented point is at **X** at the beginning, its displacement is **D** and its end location is $\mathbf{X}'$.

energy of γ-radiation) .

2. Breaks occur homogeneously, i.e. no part of the DNA strand has a higher or lower risk for break events. This hypothesis is disputable, since the DNA in the cell nucleus has spatially a very complicated crystalline structure and parts of DNA deeper inside this structure may be less exposed to radiation.

In our setting it is sensible to make the following additional assumptions.

3. Breaks are rare compared to the number of unbroken sites.

   From literature, we know that depending on experimental conditions 1 Gy radiation intensity causes on average one single strand break (ssb) in several ten thousand base pairs (bp). For example, [24] gives a value of $5.98 * 10^{-8}\mathrm{Gy}^{-1}\mathrm{Da}^{-1}$, corresponding approximately to one ssb for 25000 bp (with 660 Ga as average molecular weight of a bp). In our case, a 3.5 Gy γ-source has been used leading to a rough estimate of one ssb for 7000 bp.

Further, the considered mouse chromosomes exceed by far 10 Mbp and thus we state that

4. the total number of breaks occurring in the radiated chromosomes is large.

We are now looking for a stochastic model which suits the above assumptions. The locations of breaks constitute again a point process, now in $\mathbb{N}$. The assumed independence of breakage events, occurring at indistinguishable sites, yields a geometric distribution of fragment lengths. Moreover, we get for any (finite) subset of sites a binomially distributed number of breakages.

Due to the large numbers of bps involved in the comet experiment the discrete model is not feasible. Therefore, we want to approximate the discrete situation with some continuous pendant. The following lemma gives a hint how to do so. It states weak convergence of our discrete distributions after suitable scaling. The proof runs along well known lines.

Proposition 1 *Let $(q_k)_{k\in\mathbb{N}}$ be a sequence with $0 < q_k < 1$ and $\lim_{k\to\infty} -k \ln q_k = \lambda$.*

If $(N_k)_{k\in\mathbb{N}}$ are geometrically distributed random variables with survival probabilities $(q_k)_{k\in\mathbb{N}}$, then

$$\mathcal{L}(\frac{N_k}{k}) \xrightarrow[k\to\infty]{} \mathrm{Exp}_\lambda.$$

If $(\Xi_k)_{k\in\mathbb{N}}$ are simple point processes on $\mathbb{N}$ having for each (fixed) subset $S \subset \mathbb{N}$ with cardinality $|S| < \infty$ the probability $P(\{n_1,\ldots,n_l\} = S\cap\Xi_k) = (1-q_k)^l q_k^{|S|-l}$, then

$$\mathcal{L}(\Xi_k/k) \xrightarrow[k\to\infty]{} \Pi_\lambda,$$

where Π_λ is the stationary Poisson process on $\mathbb{R}_+$ with intensity λ and $\Xi/r = \{x/r : x \in \Xi\}$.

Proof. For the first statement observe that the characteristic function of N_k/k is

$$\hat{\mu}_k(x) = \frac{1-e^{-\lambda/k}}{1-e^{(ix-\lambda)/k}},$$

and converges for $k \to \infty$ pointwise towards the characteristic function $\lambda(\lambda - ix)^{-1}$ of an exponentially distributed random variable. The continuity theorem asserts the weak convergence of the respective probability distributions.

The second statement of the lemma is shown with Laplace functionals. The Laplace functional of a random measure M on $\mathbb{R}_+$ is defined for nonnegative functions f as $L_M(f) = \mathbb{E}(\exp(-\int_{\mathbb{R}_+} f(x)dM(x)))$. We consider Ξ_k/k as point process on $\mathbb{R}_+$ equivalently described by the random measure $\sum_{i\in\Xi_k}\delta_{i/k}$. We evaluate its Laplace functional L_k for simple functions

$$f(x) = \sum_{m=1}^{N} a_m \cdot 1_{B(m)}(x),$$

where a_i are positive numbers and $B(i)$ are finite intervals in $\mathbb{R}_+$.

With the independence assumption we get

$$\begin{aligned} L_k(f) &= \mathbb{E}(\exp(-\int_{\mathbb{R}_+}\sum_{m=1}^{N} a_m \cdot 1_{B(m)}(x)d(\frac{\Xi_k}{k}(x)))) \\ &= \mathbb{E}(\prod_{m=1}^{N}\exp(-\sum_{B(m)\cap\Xi_k/k} a_m)) \end{aligned}$$

$$
\begin{aligned}
&= \prod_{m=1}^{N} \mathbb{E}\Big(\prod_{n \in k \cdot B(m)} B(m) \cap \Xi_k / k \Big) \\
&= \prod_{m=1}^{N} \prod_{n \in k \cdot B(m)} (e^{-a_m}(1-q_k)+q_k) \\
&= \prod_{m=1}^{N} \Big(1+\frac{(\lambda+o(1))(1-e^{-a_m})}{k}\Big)^{k\cdot|B(m)|}.
\end{aligned}
$$

In the last line we use the expansion $1-\exp(\lambda/k) = \lambda/k + o(1/k)$ for $k \to \infty$.

Thus, one has for simple functions f the convergence of the Laplace functional of Ξ_k/k towards the Laplace functional of Π_λ:

$$
\lim_{k\to\infty} L_k(f) = \prod_{m=1}^{N} e^{(1-e^{-a_m})\lambda|B(m)|} = \exp\Big(\int_{\mathbb{R}_+} (1-e^{-f(x)})\lambda dx\Big).
$$

The convergence of Laplace functionals on intervals extends to all sets of the semiring generated by intervals. This implies that the finite dimensional distributions of Ξ_k/k given by

$$
P_{\Xi_k/k}(k_1,\ldots,k_m) = P\left(\sharp\left\{\frac{\Xi_k}{k}\bigcap A_1\right\} = k_1,\ldots,\sharp\left\{\frac{\Xi_k}{k}\bigcap A_m\right\} = k_m\right)
$$

converge weakly to those of Π_λ. ([11], Cor.9.1.VIII) now finishes the proof.

□

The lemma tells us that for all sufficiently long pieces of DNA the number of breaks may be be regarded as Poisson distributed with parameter $\nu_L|I|$, where ν_L is some strictly positive real constant and $|I|$ is the length of the piece.

Further, the assumed geometric distribution of fragment length may be approximated with an exponential distribution. Moreover the interval theorem ([20],4.1) states that the exponential distribution of fragment length has the same parameter ν_L as the poisson process giving the break sites.

We assume homogeneous distribution of nucleic acids along the DNA. Therefore, we use a linear correspondence of lengths and masses of fragments. In the following we exclusively consider mass distributions. Thus, we assume that the mass of a DNA fragment has a density of the form

$$
f_M(m) = \nu_M \exp(-\nu_M m), \qquad (m \geq 0).
$$

The formula suggests that we only have to know the constant ν_M in order to know the distribution of mass completely. Viewing the literature, one finds tables giving breakage rates

of DNA strings for different intensities of irradiation. These numbers could in principle be used to fix the parameter of the exponential distribution. However, the recorded results are highly dependent on experimental conditions, which unfortunately do not match our case. Therefore, we need to determine the parameter ν_M from the COMET-assay itself .

Things get more complicated if we consider the repair mechanism which controls the data for the “repair” group. We assume that:

1. different breaks are repaired independently,

2. the repair mechanism is homogeneous (it does not depend on the site of the chromosome where the break occurred),

3. there is no difference for the cell to repair breaks between short and long fragments.

Thus, breaks are considered to be independently deleted by the repair mechanism. In the language of point processes, the process of break points is *thinned*.

The following lemma is well-known.

Proposition 2 ([11, Example 8.2(a)]) *If Z is a Poisson process with intensity measure μ then the thinned configuration Z_p, where each point of Z is deleted with probability $0 \leq p \leq 1$ is Poisson distributed with intensity measure $p\mu$.*

In our application we are interested in the repair efficacy p, which could be obtained from the ratio

$$p = \frac{\nu_M^{rep}}{\nu_M^{rad}},$$

involving the parameters of the exponential distributions for the fragment mass in repaired and irradiated cells respectively. Estimations of p are highly dependent on the RBM and the assumptions on the repair mechanism. Hence, one should look for robust substitutes of p.

5 Migration of a Single Fragment — Renewal Processes and Approximation

In this part we propose a model for the migration of DNA fragments. The model allows us to calculate the conditional distribution of displacement given the mass of a fragment.

In the literature, various migration models have already been proposed. DNA-fragments are, for instance, regarded as small balls moving in a thin net of points ([4], [9], [31]) or seen as long 'snakes' creeping between big obstacles ([1], [23], [30], [31]). In most cases the models are mainly qualitative and tailored for specific experimental conditions. Our model is an adaptation of the Ogston theory [9]. We aim to make it simple enough to allow mathematical treatment, while taking into account important features.

Let us consider each cell separately. Since in our case the agarose gel concentration is very low, the gel can be assumed to be a net of randomly distributed points. We assume the DNA fragments to be solid round balls which are not stretched and roll in the gel. The modeling can easily be generalized to the case where DNA fragments are assumed to be ellipsoids, as suggested in [4]. We admit that the assumption on the geometry is quite strong and that the issue of the shape of migrating fragments is still very controversial. This point and other concerns regarding strong assumptions have to be taken into account in refined models. But, recall that in this paper we aim at a simple model which is tractable.

Now, we assume the whole cell to be a flat cylinder (see figure 3): each fragment can move in a three-dimensional space, but in fact we will not care for the displacements along the vertical axis z, because they are negligible in comparison with the displacements induced by the electric field, which is parallel to the x-axis. For the same reason we also neglect the displacements of the fragments in the y-direction. The displacements in the y-direction and z-direction are quite complicated to understand and to model. To sum up, they can be assumed approximately as complex diffusion movements. The displacement in the x-direction depends on the mass (or length) of the fragment in a way that will be specified later. In our model we will only consider the displacements in this direction. This means that we look for the projections on the x-axis of the vectors $\mathbf{X}$, $\mathbf{D}$ and $\mathbf{X}'$, which will be simply denoted as X, D and X'.

Let us consider a fixed fragment i. When the electric field is applied, fragment i begins to migrate freely with constant and mass-independent speed (denoted v_0) in the x-direction during a period T_{i1} until it collides with an hurdle (gel fiber). Mass independent speed is justified by the fact that the force of the electric field is proportional to the polarity of the DNA which should be proportional to the mass of the fragment. After the collision fragment i needs some time (S_{i1}) to bypass the hurdle, using the shortest path (see figure 4). Then it can migrate freely again during T_{i2} until it meets the next hurdle, etc. Thus, the

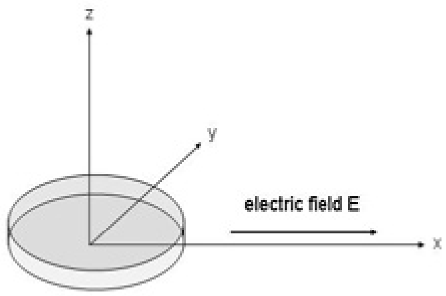


Figure 3: A whole cell in the coordinate system

migration consists of a succession of periods $T_{i1}, S_{i1}, T_{i2}, S_{i2}, \ldots, T_{ik}, S_{ik}, \ldots$ The electric field is applied at time $t = 0$ and the time t_0 corresponds to the end of the experiment, i.e. the time at which we observe the location of the fragment. The periods T_{ik} and S_{ik} can be seen as independent realizations of random variables T_k and S_k.

As a next step, we model the distribution of the $(T_k)_{k=1}^{\infty}$ and $(S_k)_{k=1}^{\infty}$.

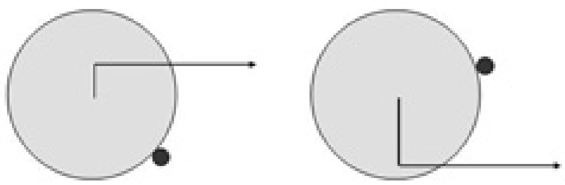


Figure 4: DNA fragments bypassing hurdles using the shorter path. The big gray disks represent DNA fragments, the small black disks represent hurdles (which are actually modelled as points). The respective paths of the fragment centers is also indicated.

5.1 The Distribution of T_k

To model T_k, we assume that the distribution of hurdles is a planar Poisson process homogeneous along the x-axis. In other words, if we follow a DNA fragment along the x-axis, we make the following assumptions.

1. We suppose the gel is perfectly homogeneous, so the probability that a fragment meets a hurdle at location between $[x, x+\Delta x]$ depends only on Δx but not on x, for all $x \geq 0$ and $\Delta x \geq 0$.

2. We also suppose that the probability for a certain fragment to bump into a hurdle is independent from where and how many times it bumped into a hurdle earlier.

3. Since the gel is very thin, we make the assumption that a fragment can not be in contact with more than one hurdle at the same time.

Under these assumptions, the number of hurdles a fragment meets on its way along the x-axis is a Poisson process with x playing the role of t. Then the distance between two hurdles is exponentially distributed with a parameter λ. We call C the number of hurdles per volume unit. To compute λ, let us imagine a round ball migrating along the x-axis. The cross section of a ball with radius a equals πa^2. Thus, during a short displacement Δx, the swept volume is $\pi a^2 \Delta x$ and the probability that the ball bumps into a hurdle is $\pi a^2 C \Delta x$, hence resulting in the simple formula $\lambda = \pi \cdot Ca^2$. Since the mass m of the fragment is roughly proportional to its volume, λ is approximately proportional to $Cm^{2/3}$.

The assumed constancy and mass-independence of the speed along the x-axis gives that the T_k are exponentially distributed random variables with the same parameter λ. The respective means can now be represented in dependence of m as $\mathbb{E}(T_k) = K_T m^{-2/3}$ for all $k > \geq 1$, with an universal constant K_T.

5.2 The Distribution of S_k

Under a quite strong assumption, the modeling of the S_k is easy. Assuming that all fragments bypass the hurdles with the same constant speed in y-direction, we get after a short computation that the S_k are uniformly distributed in the interval $[0, K_S m^{1/3}]$, with K_S being constant for all fragments.

5.3 Definition of τ

We now define for each DNA fragment and time $t \geq 0$ the integer valued random variable $\tau(t)$:

$$\tau(t) = \max\{n : \sum_{k=1}^{n-1} (T_k + S_k) < t\}. \tag{1}$$

Recall that $t = t_0$ is the time when the experiment is stopped and migration ends. The random time $\tau(t) - 1$ counts for each fragment the number of hurdles completely bypassed until t. With these settings, the displacement $D(t)$ of a given fragment along the x-axis at time t clearly satisfies

$$v_0 \cdot \sum_{k=1}^{\tau(t)-1} T_k \quad \leq \quad D(t) \quad \leq \quad v_0 \cdot \sum_{k=1}^{\tau(t)} T_k,$$

where ν_0 denotes the constant migration speed. We are interested in statements about the displacements $D := D(t_0)$ at the end of the experiment for t_0 tending to infinity.

5.4 Asymptotic of D

First we observe that, if t approaches infinity, the number of obstacles hit by a fragment goes to infinity with probability 1:

Proposition 3 *The integer valued random variables $\tau(t), t \geq 0$ satisfy*

$$P(\lim_{t\to\infty} \tau(t) = \infty) = 1.$$

Proof. Let t_k be a monotonically increasing sequence with $\lim\limits_{k\to\infty} t_k = \infty$.
We get from the definitions

$$\tau(t_n) \leq \tau(t_m) \text{ for } n < m.$$

Markov's inequality shows for all $t_k > 0$ and $N \in \mathbb{N}$

$$P(\tau(t_k) \leq N) = P(\sum_{k=1}^{N} (T_k + S_k) > t_k) \leq \frac{N \cdot \mathbb{E}(T_1 + S_1)}{t_k}.$$

Now, we define: $\quad A_N(t_k) := \{\tau(t_k) \leq N\} \quad$ and $\quad A_N := \bigcap\limits_k A_N(t_k)$.

As $\quad A_N(t_n) \supseteq A_N(t_m)$ for $n < m$, $\quad A_N(t_k)$ tends for growing k monotonically to A_N. Thus, the σ-additivity of P and the above inequality give

$$P(A_N) = \lim_{k\to\infty} P(A_N(t_k)) = 0, \text{ and further}$$

$$P(\lim_{k\to\infty} \tau(t_k) = \infty) = 1 - P(\bigcup_N A_N) \geq 1 - \sum_N P(A_N) = 1.$$

$\square$

This result allows us to establish almost sure convergence of $D(t_0)/t_0$ if the time for stopping the experiment increases to infinity.

Proposition 4

$$P\left(\lim_{t\to\infty}\frac{D(t)}{t} \;=\; v_0\,\frac{\mathbb{E}T_1}{\mathbb{E}(T_1)+\mathbb{E}(S_1)}\right)=1 \tag{2}$$

Proof. From the definitions for D and τ we immediately get ($\tau=\tau(t)$)

$$\frac{\sum_{k=1}^{\tau}T_k}{\sum_{k=1}^{\tau}(T_k+S_k)} \;\le\; \frac{D(t)}{v_0\cdot t} \;\le\; \frac{\sum_{k=1}^{\tau-1}T_k}{\sum_{k=1}^{\tau-1}(T_k+S_k)}+\frac{T_\tau}{\sum_{k=1}^{\tau-1}(T_k+S_k)}. \tag{3}$$

The strong law of large numbers for i.i.d. variables and Lemma 3 yield

$$\lim_{\tau\to\infty}\frac{1}{\tau}\sum_{k=1}^{\tau}(T_k+S_k) \;=\; \mathbb{E}(T_1+S_1) \quad \text{almost surely}, \tag{4}$$

and thus, one has the corresponding convergence in (3). To establish the required result, we observe that

$$P\left(\lim_{\tau\to\infty}\frac{T_\tau}{\sum_{k=1}^{\tau-1}(T_k+S_k)} \;=\; 0\right)=1.$$

This is valid according to (4) and the fact that $\mathbb{E}(T_1)<\infty$ is a (necessary and) sufficient condition for

$$P\left(\lim_{\tau\to\infty}\frac{T_\tau}{\tau}=0\right)=1$$

to hold ([28],lemma 7.5.1). □

The maximal displacement $D(t)=\nu_0\cdot t$ happens if a fragment meets no hurdle on its path until time t. In general $D(t)/t\le\nu_0$ holds. Therefore, the theorem on dominated convergence yields with Lemma 4

$$\lim_{t\to\infty}\frac{\mathbb{E}(D(t))}{t} \;=\; v_0\,\frac{\mathbb{E}T_1}{\mathbb{E}(T_1)+\mathbb{E}(S_1)}.$$

These results indicate that for t_0 large with respect to the order of S_k+T_k one has the reasonable approximation for $D=D(t_0)$

$$\begin{aligned} D &\approx t_0\cdot v_0\cdot\frac{\mathbb{E}T_1}{\mathbb{E}T_1+\mathbb{E}S_1} \\ &= t_0\cdot v_0\cdot\frac{K_T m^{-2/3}}{K_T m^{-2/3}+K_S m^{1/3}} \\ &= \frac{1}{K_1+K_2 m}, \end{aligned} \tag{5}$$

with suitable constants K_1 and K_2. This formula is considered in some references as the best empirical approximation of the (mean) displacement as function of the mass (or the length) [33, 30].

Many other formulas have been proposed in the literature [23, 1, 4]. These formulas are suited to model specific experimental conditions (strength of the electric field, approximate fragment size, gel properties, etc). However, the formula $D = 1/(K_1 + K_2 m)$ and slight modifications seem to prevail in applications. Observe that no other details about T_1 and S_1 than $\mathbb{E}S_1/\mathbb{E}T_1 \sim m$ enter the computation of D. If one derives similar dependencies for the movement of DNA with other models this directly leads to corresponding approximations for D.

6 Migration of the Fragment Population

To describe the migration of all fragments together we use the above introduced point process notation. We assume that the electric field generates a stochastic displacement of the fragments, i.e., each fragment moves independently from the others and from the interaction of the others with the gel. This implies that the random variables D_i are independent and, conditioned on $m_i = m$, identically distributed for every $m > 0$. Further, D_i should be independent of m_j, for $j \neq i$. Now we show that these assumptions allow a simple formulation of the migration problem involving a convolution product.

Our basic assumption for the application of the images is that the intensity in one pixel is proportional to the mass of DNA concentrated there. So, we have to consider the mass distribution for the DNA. In point process language, we consider intensity measures of the process. The mass intensity measure μ_Ξ for a random point configuration Ξ is defined as

$$\mu_\Xi(A) \;=\; \mathbb{E}\Big(\sum_{(X,m)\in\Xi} m 1_A(X)\Big)$$

for each Borel set $A \subseteq \mathbb{R}^2$.

In our situation there are two intensity measures: the start intensity μ_X and the end intensity $\mu_{X'}$.

The following assumptions now govern our migration model.

1. The DNA-breaking rate and the DNA-repairing rate are spatially homogeneous. This implies especially that X_i and m_i are independent.

2. The distribution of the displacement D_i of fragment i depends only on its mass m_i and not on X_i. There may be doubts whether this assumption is justified. Indeed, especially when the DNA-concentration is high, fragments may be broken by other

fragments during the migration. However, to get a feasible model, we assume that this effect does not play an important role.

We use f_M to denote the density of mass, i.e. $\int_{m_1}^{m_2} f_M(m) \cdot \mathrm{d}m$ is the fraction of fragments with mass between m_1 and m_2. Further, f_X denotes the start density, i.e. $\int_{x_1}^{x_2} f_X(x) \cdot \mathrm{d}x$ is the fraction of fragments between $x = x_1$ and $x = x_2$ at the beginning of the electrophoresis, or, to be more precise, the fraction of fragments whose gravity center is between $x = x_1$ and $x = x_2$. Similarly, we define the conditional densities $f_{X'|m}$ and $f_{D|m}$. $\int_{x'_1}^{x'_2} f_{X'|m}(x) \cdot \mathrm{d}x$ is the fraction of fragments between x'_1 and x'_2 at the end of the electrophoresis given the mass m, and $\int_{d_1}^{d_2} f_{D|m}(d) \cdot \mathrm{d}d$ is the fraction of fragments with displacement between d_1 and d_2 given the mass m. Further, let f_{μ_X} denote the density of μ_X and $f_{\mu_{X'}}$ the density of $\mu_{X'}$.

Because of assumption 1, $f_X = f_{\mu_X}$. Finally, let M_c denote the total mass of DNA contained in the considered cell c. Our model leads us to a convolution problem.

Proposition 5 *With the global density of displacement*

$$f_\sigma(d) \;=\; \frac{1}{M_c}\int_0^\infty m \cdot f_M(m) \cdot (f_{D|m})(d) \cdot \mathrm{d}m,$$

we have

$$f_{\mu_{X'}} = f_{\mu_X} * f_\sigma.$$

Proof.

The density $f_{\mu_{X'}}$ of $\mu_{X'}$ can be written as

$$f_{\mu_{X'}}(x') \;=\; \frac{1}{M_c}\int m \cdot f_M(m) \cdot f_{X'|m}(x')\mathrm{d}m.$$

From

$$f_{X'|m}(x') \;=\; \int f_X(x) \cdot f_{D|m}(x'-x)\mathrm{d}x$$

we immediately find our assertion:

$$\begin{aligned} f_{\mu_{X'}}(x') \;&=\; \frac{1}{M_c}\int f_X(x) \cdot \int m \cdot f_M(m) \cdot f_{D|m}(x'-x)\mathrm{d}x\,\mathrm{d}m \\ &=\; \int f_X(x) \cdot f_\sigma(x'-x)\mathrm{d}x. \end{aligned}$$

□

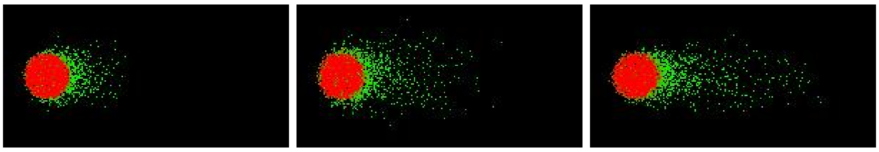


Figure 5: Simulations from the point process model for different parameters

7 Simulation and Comparison to Data

In this section, a simulation is carried out to check qualitatively whether the model for the DNA mass distribution and the model for the DNA migration lead to comet-like shapes. Subsequently, we present a simple method which allows a rough estimate of the model parameters from two specific histograms. These histograms represent, respectively, the horizontal distribution at the beginning and at the end of the electrophoresis. One of these parameters, ν_M, determines the (exponential) distribution of fragment masses.

7.1 Simulation of the DNA migration

With the software package AntsInFields [15] we implemented the above model, leaving aside the problem of calculating suitable variances. The length of the fragments was sampled from an exponential distribution. The number of fragments was fixed beforehand and assumed to be uniformly distributed over a ball. The distribution of the displacement D was taken as bivariate normal with expectation $(\frac{1}{K_1+K_2m}, 0)$, to take into account randomness of D. The variances σ_x and σ_y were fixed independently of the fragment length m and covariance was assumed to be 0. We stopped the simulations after suitable times to find comet-like shapes.

As Figure 5 indicates, the model is able to capture at least the comet-like shape of the real-world data. The program written in Oberon is available on request from the last author.

7.2 A simple method to estimate the model parameters

The model presented above includes two steps of modeling:

1. the modeling of the distribution of masses as exponential with parameter ν_M:

$$f_M(m) = \nu_M \exp(-\nu_M m),$$

2. the modeling of the dependency between the displacement D and the mass m. The problem of determination of a correct variance formula is ignored. For simplicity the displacement given the mass is assumed to equal its mean:
$$D(m) = \frac{1}{K_1 + K_2 m}.$$

The representation of $f_\sigma(d)$ according to Lemma 5 yields with setting $M_c = \frac{1}{\nu_M}$, collapsing $f_{D|m}(d') = \delta_{d'D(m)}$ and changing of variables to $m = \frac{1}{K_2}(\frac{1}{d'} - K_1)$

$$f_\sigma(d) = \frac{1}{M_c}\int_0^\infty \frac{\nu_M}{K_2 d'^2}(\frac{1}{d'} - K_1)\exp(-\frac{\nu_M}{K_2}(\frac{1}{d'} - K_1)) \quad \delta_{d'd} \quad \mathrm{d}d' \tag{6}$$
$$= \frac{\nu_M^2}{d^2 K_2^2}(\frac{1}{d} - K_1)\exp(-\frac{\nu_M}{K_2}(\frac{1}{d} - K_1)). \tag{7}$$

Although this model involves 3 parameters (ν_M, K_1 and K_2), it has only two degrees of freedom, since K_2 and ν_M both scale m and therefore appear only in the ratio $\frac{K_2}{\nu_M}$. Consequently, we can only identify the two parameters K_1 and $K = \frac{\nu_M}{K_2}$. Note that this poses no problem to the estimation of the repair efficacy p since K_2 does only depend on the gel and field properties and thus cancels in the ratio $p = \frac{\nu_M^{rep}}{\nu_M^{rad}} = \frac{K^{rep}}{K^{rad}}$.

We now concentrate on the estimation of K (and K_1). Let us consider two images from the same mouse: an image of a control cell and an image of a degraded cell. Using a JAVA program, we sum the intensities of all the pixel columns successively, for both images. Thus we obtain discretized estimates of f_{μ_X} and f_{μ_Y}, as depicted in figure 6. Notice that we have aligned the two images arbitrarily. As will become clear later, this causes no problem. To estimate the parameters K_1 (whose unit is $pixel^{-1}$) and K, we proceed as follows. For different values of K_1 and K, we perform a discrete convolution of $\hat{f}_{\mu_X}$ and f_σ. Our goal is to find the values for which a certain dissimilarity function between this convolution product and the observed $\hat{f}_{\mu_Y}$ is minimal. Since we do not know the location of the axis origin in the histogram $\hat{f}_{\mu_Y}$, this dissimilarity measure has to be translation invariant. A simple method is to match $\hat{f}_{\mu_X} * f_\sigma$ with the histogram $\hat{f}_{\mu_Y}$ using a criterion of minimal quadratic transportation costs which is an L^2-Wasserstein metric ([26],1992). Namely let the stochastic matrix $q(i,j)$ be the (unique) minimum of

$$q \longrightarrow \sum_{i,j} \hat{f}_{\mu_X} * f_\sigma(i) q(i,j)(j-i)^2$$

among all q with

$$\sum_i \hat{f}_{\mu_X} * f_\sigma(i) q(i,j) = \hat{f}_{\mu_Y}(j) \text{ for all j},$$

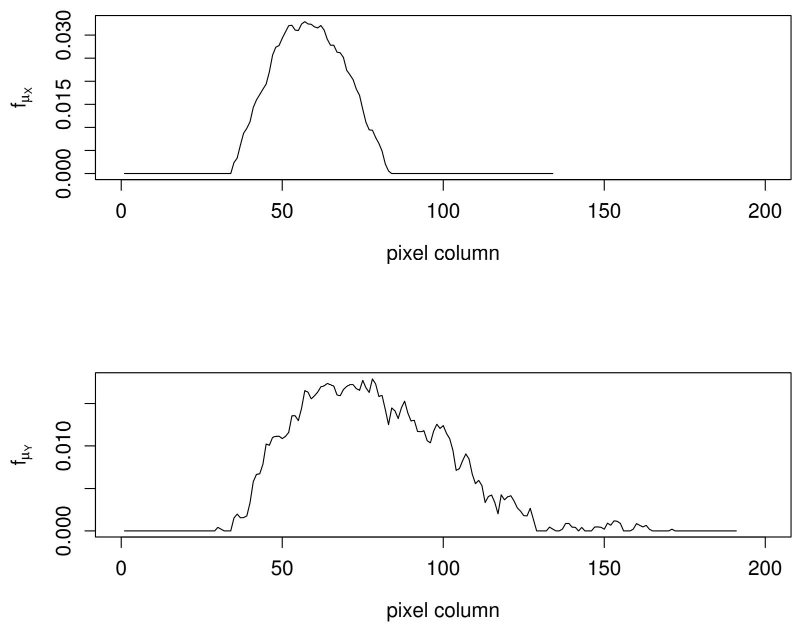


Figure 6: Histograms of a control cell (top) and a damaged cell (bottom) from the same mouse

as described in [7]. Then we take the variance

$$V = \sum_{i,j} q(i,j)(j-i)^2 - \left(\sum_{i,j} q(i,j)(j-i)\right)^2$$

of the resulting histogram as dissimilarity measure. Note, that this measure is translation invariant and it is higher for 'very different' histograms than for 'similar' histograms.

To minimize this criterion, we employ the R program `optim` which implements the optimization method of [8] and allows to give as inputs lower and upper bounds for each parameter. Here, we set the lower bounds to zero, because the parameters K and K_1 have to be strictly positive. This method yields estimates for K and K_1.

7.3 Results of the parameter estimation

For the two histograms depicted in Figure 6, the optimization algorithm yields the following parameter estimates:

$$\hat{K_1} = 0.0074$$

$$\hat{K} = 74.$$

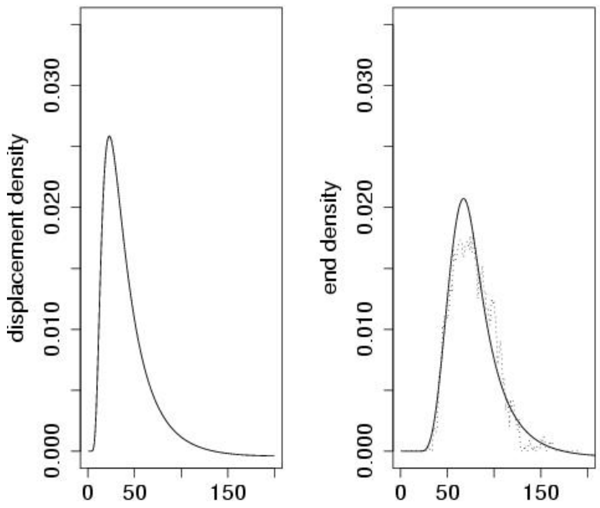


Figure 7: Histogram of the estimated displacement density (left) and estimated end density (right, solid) for $\hat{K}_1 = 0.0074$ and $\hat{K} = 74$. On the right panel, the observed end density of the damaged cell is drawn for comparison (dotted).

For these values, the displacement density $\hat{f}_\sigma$ is depicted in Figure 7 (left). To evaluate the quality of the estimation, we superpose the result of the convolution product of $\hat{f}_{\mu_X} * \hat{f}_\sigma$ obtained with the estimated parameters and the observed $\hat{f}_{\mu_Y}$, as depicted in figure 7 (right). The estimate fits the data well, which indicates that our model is quite realistic.

8 Discussion

In this work, we introduce a stochastic model to describe the COMET-assay experiment. The model comprises two parts. The first part, known in the literature as 'Random Breakage Model' deals with the distribution of length of the DNA fragments. The second part describes the migration of DNA fragments among gel fibers.

The discussed model strongly simplifies the complex mechanisms of DNA damage and electrophoresis. However, it allows mathematical analysis of the obtained cell images, a great advantage compared to more complicated (and unfeasible) models. Moreover, simulations show that the model captures phenomenological aspects quite well. A simple approach to estimate the model parameters is presented in section 7. It is a feature

of this model that only two parameters $\hat{K}_1$ and $\hat{K}$ are identifiable from a COMET-assay experiment. Therefore, an additional gauge experiment, measuring the mass dependence of migration under COMET-assay conditions, would be required to get estimates for the scale parameters ν_M and K_2.

Further, a robust matching method is requested to use our estimation for 30 control cells and 30 damaged cells which are available for each mouse. Hence, the robustness of the proposed method has to be improved. Since it is not clear whether all cells of the same mouse are equally damaged, the study of robustness might be quite difficult. However, we think that our model is able to extract relevant and easily interpretable information about damage and repair mechanisms in DNA which are not available from the classical geometric parameters.

Acknowledgements We thank C. Schindewolf for introduction into the problem, J. Georgii for the development of a JAVA platform, and F.Friedrich for the simulation via the AntsinFields platform [15]. For valuable input we also thank R. Anderson, T. Speed and F. Eckhardt-Schupp. This work was supported by the DFG, SFB 386 ”Statistical Analysis of discrete Structures” and BMBF-”Bioinformatics for the Functional Analysis of Mammalian Genomes”.

References

[1] B. Åkerman, Cyclic migration of DNA in gels: DNA stretching and electrophoretic mobility, *Electrophoresis*, 17, 1027-1036(1996).

[2] G. Alsmeyer, *Erneuerungstheorie*, Teubner, Stuttgart, 1991.

[3] J. Ashby, H. Tinwell, P.A. Lefevre and M.A. Browne, The single cell gel electrophoresis assay for induced DNA damage (COMET-assay): measurement of tail length and moment, *Mutagenesis*, 10, 85-90(1995).

[4] J.C. Bearden, Electrophoretic mobility of high-molecular-weight double-stranded DNA on agarose gels, *Gene*, 6, 221–234(1979).

[5] R.N. Bhattacharya and E.C. Waymire, *Stochastic processes with applications*, Wiley series in probability and mathematical statistics, Chichester, 1990.

[6] W. Böcker, T. Bauch, W.U. Müller and C. Streffer, Image analysis of COMET-assay measurements, *Int.J.Radiat.Biol*, 72, 449–460(1997).

[7] A.-L. Boulesteix, V. Hösel and V. Liebscher, A comparative study of empirical deconvolution techniques with application to the COMET-assay. In preparation.

[8] R.H. Byrd, P. Lu, J. Nocedal and C. Zhu, A limited memory algorithm for bound constrained optimization, *SIAM J.Sci. Comput.*, 16, 1190–1208(1995).

[9] C.R. Calladine, C.M.Collis, H.R.Drew and M.R.Mott, A study of electrophoretic mobility of DNA in agarose and polyacrylamide gels, *J.Mol.Biol.*, 221, 981–1005(1991).

[10] A.N. Shiryayev, *Probability*, Graduate Texts in Mathematics, Springer-Verlag, New York, 1984.

[11] D. Daley and D. Vere-Jones, *An Introduction to the Theory of Point Processes*, Springer-Verlag, New York, 1988.

[12] M. De Boeck, N. Touil, G. De Visscher, P. Aka Vande and M. Kirsch-Volders, Validation and implementation of an internal standard in comet assay analysis, *Mutat.Res.*, 469, 181–197(2000).

[13] D.W. Fairbairn, P.L. Olive and K.L. O'Neill, The COMET-assay: a comprehensive review, *Mutat.Res.*, 339, 37–59(1995).

[14] W.Feller, *An introduction to probability theory and its application, Vol.II*, Wiley series in probability and statistics, Wiley, Chichester, 1971.

[15] F. Friedrich, The software package AntsInFields, `http://www.antsinfields.de`, 2002.

[16] G. Grimmett and D. Stirzaker, *Probability and random processes*, Oxford science publications, Oxford,1992.

[17] Introduction to COMET assay, `http://www.cometassay.com`.

[18] A.F. Karr, *Point processes and their statistical inference*, Probability: pure and applied, Manuel Neuts, Paris, 1986.

[19] C.R.H. Kent, J.J. Eady, G.M. Ross and G.G. Steel, The comet moment as a measure of DNA damage in the COMET-assay, *Int.J.Radiat.Biol.*, 67, 655–660(1995).

[20] J.F.C. Kingman, *Poisson processes*, Oxford Studies in Probability, Clarendon Press, Oxford, 1993.

[21] F. Kraxenberger, K.J. Weber, A.A. Friedl, F. Eckardt-Schupp, M. Flentje, P. Quicken and A.M. Kellerer, DNA double-strand breaks in mammalian cells exposed to γ-rays and very heavy ions, *Radiat.Environ.Biophys.*, 37, 107–115(1998).

[22] P.R. Krishnaiah and P.K. Sen, *Handbook of statistics 4*, Elsevier Science Publishers, Amsterdam,1984.

[23] M. Lalande, J. Noolandi, C. Turmel, R. Brousseau, J. Rousseau and G.W. Slater, Scrambling of bands in gel electrophoresis of DNA, *Nucleic Acids.Res.*, 16, 5427–5437(1988).

[24] J.R. Milligan, J.A. Aguillera, R.A. Paglinawan, J.F. Ward and C.L. Limoli, DNA strand break yields after post-high LET irradiation incubation incubation with endonuclease-III and evidence for hydroxyl radical clustering, *Int. J. Radiat. Biol.*, 77,2,155-164(2001).

[25] R. Nelson, *Probability, stochastic processes and queuing theory*, Springer-Verlag, Berlin,1995.

[26] S.T. Rachev, *Probability metrics and the stability of stochastic models*, Wiley series in probability and mathematical statistics, Chichester, 1991.

[27] S.I. Resnick, *Adventures in Stochastic Processes*, Birkhuser, Berlin, 1994.

[28] S.I. Resnick, *A probability path*, Birkhuser, Berlin, 1999.

[29] S.M. Ross, *Stochastic processes*, Wiley series in probability and mathematical statistics, Chichester,1983.

[30] P. Serwer, Sieving of double-stranded DNA during agarose gel electrophoresis, *Electrophoresis*, 10, 327–331(1989).

[31] G. W. Slater, P.Mayer and G.Drouin, Migration of DNA through gels, *Methods Enzymol.*, 270, 272–295(1996).

[32] Software for Comet Analysis: VISCOMET, Manual TILL Photonics

[33] E.M. Southern, Measurement of DNA Length by Gel Electrophoresis, *Anal.Biochem.*, 100, 319–323(1979).

JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS,VOL.5,NO.1,77-88,2007,COPYRIGHT 2007 EUDOXUS PRESS,LLC

Formulae of numerical differentiation

Maxim Dvornikov
Institute of Terrestrial Magnetism, Ionosphere and Radiowave Propagation (IZMIRAN)
142190, Troitsk, Moscow region, Russia
E-mail: maxdvo@izmiran.ru

Abstract

We derived the formulae of central differentiation for the finding of the first and second derivatives of functions given in discrete points, with the number of points being arbitrary. The obtained formulae for the derivative calculation do not require direct construction of the interpolating polynomial. As an example of the use of the developed method we calculated the first derivative of the function having known analytical value of the derivative. The result was examined in the limiting case of infinite number of points. We studied the spectral characteristics of the weight coefficients sequence of the numerical differentiation formulae. The performed investigation enabled one to analyze the accuracy of the numerical differentiation carried out with the use of the developed technique.

Mathematics Subject Classification: Primary 65D25; Secondary 65T50

1 Introduction

In solving many mathematical and physical problems by means of numerical methods one is often challenged to seek derivatives of various functions given in discrete points. In such cases, when it is difficult or impossible to take derivative of a function analytically one resorts to numerical differentiation.

It should be noted that there exists a great deal of formulae and techniques of numerical differentiation (see, for instance, Ref. [1]). As a rule, the function in question $f(x)$ is replaced with the easy-to-calculate function $\varphi(x)$ and then it is approximately supposed that $f'(x) \approx \varphi'(x)$. The derivatives of higher orders are computed similarly. Therefore, in order to obtain numerical value of the derivative of the considered function it is necessary to indicate correctly the interpolating function $\varphi(x)$. If the values of the function $f(x)$ are known in s discrete points, the function $\varphi(x)$ is usually taken as the polynomial of $(s-1)$th power.

To find the derivative of functions having the intervals both quick and slow variation quasi-uniform nets are used (see Ref. [2]). This method has an ad-

vantage since constant small mesh width is unfavorable in this case, because it leads to the strong enhancement of the function values table.

The problem of the numerical differentiation accuracy is also of interest. The numerical differentiations formulae, taking into account the values of the considered function both at $x > x_0$ and $x < x_0$ (x_0 is a point where the derivative is computed), are called central differentiation formulae. For instance, the formulae based on Stirling interpolating polynomial can be included in this class. Such formulae are known to have higher accuracy compared to the formulae, using unilateral values of a function[1], i.e., for instance, at $x > x_0$.

The range of numerical differentiation formulae based on different interpolating polynomials is limited, as a rule, to finite points of interpolation. All available formulae known at the present moment are obtained for a certain concrete limited number of interpolation points (see Refs. [3, 4]). It can be explained by the fact that the procedure of the finding of the interpolating polynomial coefficients in the case of the arbitrary number of interpolation points is quite awkward and requires formidable calculations.

It is worth mentioning that the procedure of the numerical differentiation is incorrect. Indeed, in Ref. [2] it was shown that it is possible to select such decreasing error of the function in question which results in the unlimited growth of the error in its first derivative.

Some recent publications devoted to the numerical differentiation problem should be mentioned (see, e.g., Ref. [5]). In this work the finite difference formulae for real functions on one dimensional grids with arbitrary spacing were considered.

The formulae of central differentiation for the finding of the first and the second derivatives of the functions given in $(2n + 1)$ discrete points are derived in this paper. The number of interpolation points is taken to be arbitrary. The obtained formulae for the derivatives calculation do not require direct construction of the interpolating polynomial. As an example of the use of the developed method we calculate the first derivative of the function $y(x) = \sin x$. The obtained result is studied in the limiting case $n \to \infty$. We examine the spectral characteristics of the weight coefficients sequence of the numerical differentiation formulae for the different number of the interpolation points. The performed analysis can be applied to the studying of the accuracy of the numerical differentiation technique developed in this work. It is found that the derived formulae of numerical differentiation have a high accuracy in a very wide range of spatial frequencies.

[1] The formulae using, for example, Newton interpolating polynomial are attributed to this class of numerical differentiation formulae.

2 Formulae for approximate values of the first and the second derivatives

Without the restriction of generality we suppose that the derivative is taken in the zero point, i.e. $x_0 = 0$. Let us consider the function $f(x)$ given in equidistant points $x_m = \pm mh$, where $m = 0, \ldots, n$ and h is the constant value. We can pass the interpolating polynomial of the $2n$th power through these points

$$P_{2n}(x) = \sum_{k=0}^{2n} c_k x^k, \tag{2.1}$$

the values of the function in points of interpolation $f_m = f(x_m)$ coinciding with the values of the interpolating polynomial in these points: $P_{2n}(x_m) = f_m$. Let us define as d_m the differences of the values of the function $f(x)$ in diametrically opposite points x_m and x_{-m}, i.e. $d_m = f_m - f_{-m}$. We can present d_m in the form

$$d_m = 2\sum_{k=0}^{n-1} c_{2k+1} h^{2k+1} m^{2k+1}. \tag{2.2}$$

To find the coefficients c_{2k+1}, $k = 0, \ldots n-1$, we have gotten the system of inhomogeneous linear equations with the given free terms d_m. It will be shown below that this system has the single solution.

We will seek the solution of the system [Eq. (2.2)] in the following way

$$c_{2k+1} = \frac{1}{2h^{2k+1}} \sum_{m=1}^{n} d_m \alpha_m^{(2k+1)}(n), \tag{2.3}$$

where $\alpha_m^{(2k+1)}(n)$ are the undetermined coefficients satisfying the condition

$$\sum_{m=1}^{n} \alpha_m^{(2l+1)}(n) m^{2k+1} = \delta_{lk}, \quad l, k = 0, \ldots, n-1. \tag{2.4}$$

Thus, the system of equations [Eq. (2.2)] is reduced to the equivalent, but more simple system [Eq. (2.4)], in which for each fixed number $k = 0, \ldots, n-1$ it is necessary to find the coefficients $\alpha_m^{(2l+1)}(n)$.

Let us resolve the system of equations [Eq. (2.4)] according to the Cramer's rule:

$$\alpha_m^{(2l+1)}(n) = \frac{\Delta_m^{(2l+1)}(n)}{\Delta_0(n)}, \tag{2.5}$$

where

$$\Delta_0(n) = \begin{vmatrix} 1 & 2 & \ldots & n \\ 1 & 2^3 & \ldots & n^3 \\ \ldots & \ldots & \ldots & \ldots \\ 1 & 2^{2n-1} & \ldots & n^{2n-1} \end{vmatrix} = n! \prod_{1 \le i < j \le n} (j^2 - i^2) \neq 0, \tag{2.6}$$

$$\Delta_m^{(2l+1)}(n) = \begin{vmatrix} 1 & 2 & \dots & m-1 & 0 & m+1 & \dots & n \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & 2^{2l+1} & \dots & (m-1)^{2l+1} & 1 & (m+1)^{2l+1} & \dots & n^{2l+1} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & 2^{2n-1} & \dots & (m-1)^{2n-1} & 0 & (m+1)^{2n-1} & \dots & n^{2n-1} \end{vmatrix}.$$

In Eq. (2.6) we used the formula for the calculation of the Vandermonde determinant. From Eq. (2.6) it follows that the determinant of the system of equations [Eq. (2.4)] is not equal to zero, i.e. the system of equations [Eq. (2.2)] has the single solution.

The most simple expression for $\Delta_m^{(2l+1)}(n)$ is obtained in the case of $l = 0$ that corresponds to a calculation of the first-order derivative

$$\Delta_m^{(1)}(n) = (-1)^{m+1} \left(\frac{n!}{m}\right)^3 \prod_{\substack{1 \le i < j \le n \\ i,j \ne m}} (j^2 - i^2). \tag{2.7}$$

From Eq. (2.5) as well as taking into account Eqs. (2.6) and (2.7) we get the expression for the coefficients $\alpha_m^{(1)}(n)$

$$\alpha_m^{(1)}(n) = \frac{1}{m\pi_m(n)}, \tag{2.8}$$

where

$$\pi_m(n) = \prod_{\substack{k=1 \\ k \ne m}}^{n} \left(1 - \frac{m^2}{k^2}\right). \tag{2.9}$$

It should be noted that one can similarly get the expression for the coefficients $\alpha_m^{(2n-1)}(n)$ which is presented in the following way

$$\alpha_m^{(2n-1)}(n) = \frac{(-1)^{n+1} m}{(n!)^2 \pi_m(n)}.$$

Taking into account Eqs. (2.1)-(2.3) we finally get the formula for the first derivative of the function $f(x)$

$$f'(0) \approx P'_{2n}(0) = \frac{1}{2h} \sum_{m=1}^{n} \alpha_m^{(1)}(n)(f_m - f_{-m}). \tag{2.10}$$

The algorithm for the computation of the coefficients $\alpha_m^{(1)}(n)$ is presented in the appendix A and the results for the certain concrete number of the interpolation points $(2n + 1)$ are given in Tab. 1. Note that the expression for the first derivative obtained by this method coincides with the value presented in the Refs. [3, 4] for $n = 1, 2$ that corresponds to three and five points of interpolation. However, technique developed in this article allows one to calculate the coefficients $\alpha_m^{(1)}(n)$, and hence the first derivative, for any value of n.

n	$\alpha_1^{(1)}(n)$	$\alpha_2^{(1)}(n)$	$\alpha_3^{(1)}(n)$	$\alpha_4^{(1)}(n)$	$\alpha_5^{(1)}(n)$	$\alpha_6^{(1)}(n)$
1	1	0	0	0	0	0
2	$4/3$	$-1/6$	0	0	0	0
3	$3/2$	$-3/10$	$1/30$	0	0	0
4	$8/5$	$-2/5$	$8/105$	$-1/140$	0	0
5	$5/3$	$-10/21$	$5/42$	$-5/252$	$1/630$	0
6	$12/7$	$-15/28$	$10/63$	$-1/28$	$2/385$	$-1/2772$

Table 1: Values of the coefficients $\alpha_m^{(1)}(n)$.

Similar formula can be obtained for the calculation of the second derivative. We give without proof corresponding expression

$$f''(0) \approx P''_{2n}(0) = \frac{1}{h^2} \sum_{m=1}^{n} \alpha_m^{(2)}(n)(f_m - 2f(0) + f_{-m}), \tag{2.11}$$

where

$$\alpha_m^{(2)}(n) = \frac{1}{m^2 \pi_m(n)}, \tag{2.12}$$

and the product $\pi_m(n)$ is introduced in Eq. (2.9).

As an example of the use of the obtained central differentiation formulae we will compute the first derivative of the function $y(x) = \sin x$ at $x = 0$. Let us set the value of the mesh width h equal to $\pi/2$. Notice that, as a rule, the less the mesh width h the more exact result numerical differentiation gives. We have chosen rather big value of h. The Eq. (2.10) for this case takes the form

$$y'(0) = \frac{2}{\pi} \sum_{m=0}^{n} (-1)^m \alpha_{2m+1}^{(1)}(n). \tag{2.13}$$

In Eq. (2.13) we take that $d_m = 0$ if m is an even number, and $d_m = \pm 2$ if m is an odd number.

Let us study the obtained result in the limiting case $n \to \infty$. First, it is necessary to calculate the value of the product $\pi_m(n)$ within the limit $n \to \infty$

$$\pi_m = \lim_{n\to\infty} \pi_m(n) = \lim_{\varepsilon\to 0} \frac{1}{1 - \left(\frac{m+\varepsilon}{m}\right)^2} \prod_{k=1}^{\infty} \left(1 - \frac{(m+\varepsilon)^2}{k^2}\right) = \frac{(-1)^{m+1}}{2} \lim_{\varepsilon\to 0} \frac{\sin \pi\varepsilon}{\pi\varepsilon} = \frac{(-1)^{m+1}}{2} \tag{2.14}$$

Here we used the known value of infinite product

$$\prod_{k=1}^{\infty} \left(1 - \frac{x^2}{k^2}\right) = \frac{\sin \pi x}{\pi x}.$$

Using Eqs. (2.8) and (2.14), we find that the expression for the coefficients $\alpha_m^{(1)}(n)$ within the limit $n \to \infty$ is represented in the following way

$$\alpha_m^{(1)} = \lim_{n\to\infty} \alpha_m^{(1)}(n) = (-1)^{m+1}\frac{2}{m}. \tag{2.15}$$

Now it is easy to complete the studying of Eq. (2.13). Substituting the result from Eq. (2.15) to Eq. (2.13) and using the known value of infinite series we get that

$$y'(0) = \frac{4}{\pi}\sum_{m=0}^{\infty}\frac{(-1)^m}{2m+1} = 1.$$

Thus, it is shown that the method of the derivatives finding, developed in this paper, gives for the function $y(x) = \sin x$ the value of the first derivative which coincides with the exact analytical one even at rather crude mesh width.

3 Spectral characteristics of weight coefficients sequences

In the section 2 of the present work we derived the formula for the finding of the first derivative of the function $f(x)$ at $x = 0$. This result can be easily generalized for the case of the arbitrary point $x = kh$. If we set that $\alpha_{-m}^{(1)}(n) = -\alpha_m^{(1)}(n)$, and moreover supposing that $\alpha_m^{(1)}(n) = 0$ for $m \equiv 0$ and $m > n$ (see Tab. 1), then in the considered case Eq. (2.10) reads as follows

$$f'(kh) \approx P'_{2n}(kh) = \frac{1}{2h}\sum_m \alpha_{m-k}^{(1)}(n) f_m. \tag{3.1}$$

Here the summing is taken over all range of the function involved: $f(kh)$. For instance, if the values of the function are set on the limited equidistant collection of elements N, then Eq. (3.1) can be rewritten in the form

$$f'(kh) \approx \frac{1}{2h}\sum_{m=0}^{N-1} \alpha_{m-k}^{(1)}(n) f_m, \quad N \geq 2n. \tag{3.2}$$

It is worth noticing that in Eq. (3.2) we used the periodicity condition of the weight coefficients

$$\alpha_m^{(1)}(n) = \alpha_{m-N}^{(1)}(n).$$

Fig. 1 presents the example of the weight coefficients of the differentiating sequence $\alpha_m^{(1)}(1)$. Thus the first derivative computation of the function $f(x)$ at the points $x = kh$, $k = 0, 1, \ldots, N-1$, is reduced to the procedure of the calculation of the mutual correlation function between the finite sequences $\alpha_m^{(1)}(n)$ and f_m.

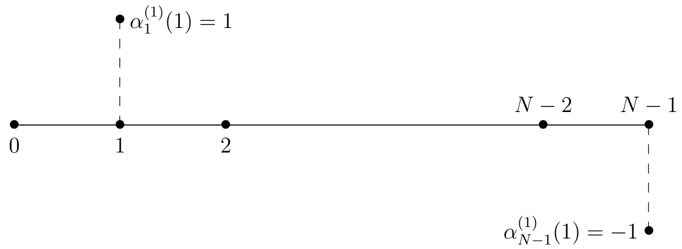


Figure 1: The coefficients $\alpha_m^{(1)}(1)$.

It is known (see, e.g., Ref. [6]) that if a function satisfies the Dirichlet conditions in the interval $(-l, l)$, then it can be expanded into the Fourier series

$$f(x) = \sum_{k=-\infty}^{+\infty} c_k \exp\left(i\frac{\pi k}{l}x\right), \tag{3.3}$$

where the expansion coefficients are presented in the way

$$c_k = c_{-k}^* = \frac{1}{2l}\int_{-l}^{l} f(\xi)\exp\left(-i\frac{k\pi}{l}\xi\right)d\xi.$$

If the first derivative $f'(x)$ satisfies the analogous conditions as the function $f(x)$, then the following expression will be valid

$$f'(x) = \sum_{k=-\infty}^{+\infty}\left\{i\frac{k\pi}{l}\right\} c_k \exp\left(i\frac{\pi k}{l}x\right), \tag{3.4}$$

Therefore, form Eqs. (3.3) and (3.4) it follows that the differentiation procedure is the linear filter with the frequency characteristic: $\mathfrak{K}_1(k) = ik(\pi/l)$ [7]. Similarly we receive for the second derivative

$$f''(x) = \sum_{k=-\infty}^{+\infty}\left\{-\left(\frac{k\pi}{l}\right)^2\right\} c_k \exp\left(i\frac{\pi k}{l}x\right),$$

In this case the the frequency characteristic of the corresponding filter has the form: $\mathfrak{K}_2(k) = -k^2(\pi/l)^2$.

According to Wiener-Khinchin theorem (see, e.g., Ref. [7]) the mutual correlation function between the two finite sequences can be calculated with the help of the inverse Fourier transform of the mutual spectrum of the considered sequences. Thus, if we define that

$$\beta_1(r) = \sum_{m=0}^{N-1}\alpha_m^{(1)}(n)\exp\left(-i\frac{2\pi}{N}mr\right),$$

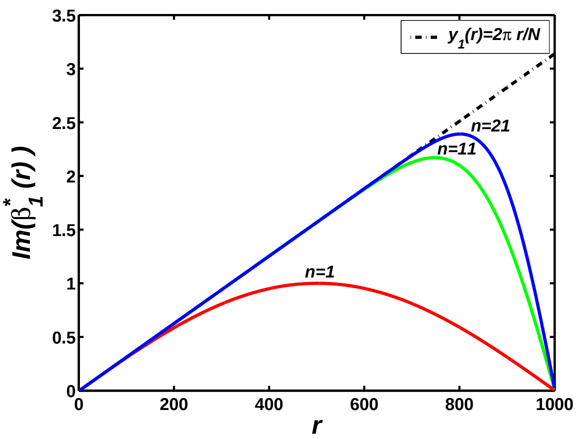


Figure 2: The spectra of various sequences $\alpha_m^{(1)}(n)$ at $N = 2000$.

is the complex spectrum of the differentiating sequence $\alpha_m^{(1)}(n)$, and

$$c_r = \sum_{m=0}^{N-1} f_m \exp\left(-i\frac{2\pi}{N}mr\right),$$

is the spectrum of the function $f(x)$, then it follows from Eq. (3.2) that

$$f'(kh) = \frac{1}{2h}\sum_{r=0}^{N-1} c_r \beta_1^*(r) \exp\left(i\frac{2\pi}{N}kr\right), \tag{3.5}$$

where $\beta_1^*(r)$ is the complex conjugated quantity with respect to $\beta_1(r)$.

Comparing Eqs. (3.4) and (3.5) we obtain that the accuracy of the numerical differentiation performed with the use of the various types of the sequences $\alpha_m^{(1)}(n)$ is characterized by the closeness of imaginary parts of their spectra to the linearly growing sequence $y_1(r) = 2\pi r/N$.

The spectra of the sequences $\alpha_m^{(1)}(n)$ are depicted in Fig. 2 for the various values of n at $N = 2000$. It can be seen from this figure that for $n = 1$, i.e. for the sequence shown in Fig. 1, the imaginary part of the spectrum is the branch of the function $\sin(2\pi r/N)$. The linearity condition is satisfied only in the vicinity of zero and $N/2$. However, at $n = 11$ the spectrum practically does not differ from the linear one up to $r \approx N/2$. The more close to linear one is the spectrum of the sequence $\alpha_m^{(1)}(21)$.

The difference between the imaginary parts of the spectra of the sequences $\alpha_m^{(1)}(n)$ and the linearly growing sequence $y_1(r) = 2\pi r/N$ are presented in Fig. 3. The computations have been performed with the accuracy up to 10^{-15}, thus the reliable results at $n = 11$ have been obtained for $r \gtrsim 150$, and at $n = 21$ for $r \gtrsim 300$. The presented results demonstrate the high accuracy of the numerical

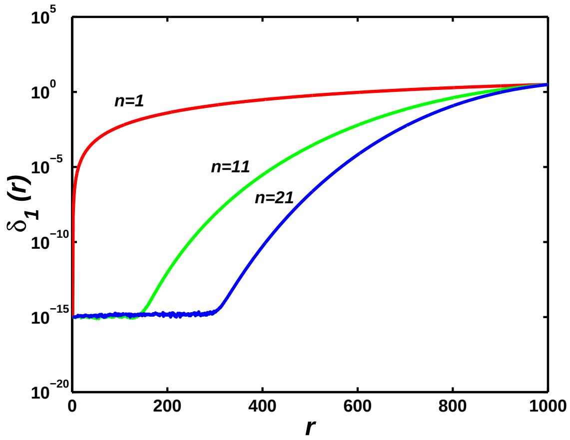


Figure 3: The function $\delta_1(r) = \Im\mathfrak{m}\big(\beta_1^*(r)\big) - y_1(r)$ versus r for different n.

differentiation carried out with the help of the sequences $\alpha_m^{(1)}(n)$ in the wide range of the spatial frequencies.

Now let us briefly consider the sequences for the calculation of the second derivative $\alpha_m^{(2)}(n)$, which are given in Eq. (2.12). Their spectral properties can be obtained in the similar manner as we have done it for the case of the sequences $\alpha_m^{(1)}(n)$ and therefore we just present the final results. The spectra of the sequences $\alpha_m^{(2)}(n)$ are shown in Fig. 4. It follows from this figure that the closeness of the corresponding spectrum to the parabola $y_2(r) = -(2\pi r/N)^2$ in the case of $n = 1$ exists only in the vicinity of zero. The spectra at $n = 11$ and $n = 21$ are close to function $y_2(r)$ in a wider range of r ($r \lesssim 750$ and $r \lesssim 800$ respectively). The difference between the real parts of the spectra of the sequences $\alpha_m^{(2)}(n)$ and the parabola $y_2(r) = -(2\pi r/N)^2$ are depicted in Fig. 5 in the logarithmic scale. This figure again demonstrates the high accuracy of the second derivative computation with the use of the sequences $\alpha_m^{(2)}(n)$.

4 Conclusion

In conclusion we note that the method of central differentiation formulae finding has been developed in this article. The elaborated technique does not require direct construction of the interpolating polynomial. We have derived simple and convenient expressions for the first and the second derivatives [Eqs. (2.10) and (2.11)] of the function given in $(2n+1)$ discrete points. The number n was taken to be arbitrary. In contrast to the results of the Ref. [5], where the recursion relations for the calculation of the weight coefficient being used in numerical differentiation formulae were considered, in the present work the expressions for the considered weight coefficients have been derived in the explicit form for the arbitrary number of interpolation points. As an example of the use of the

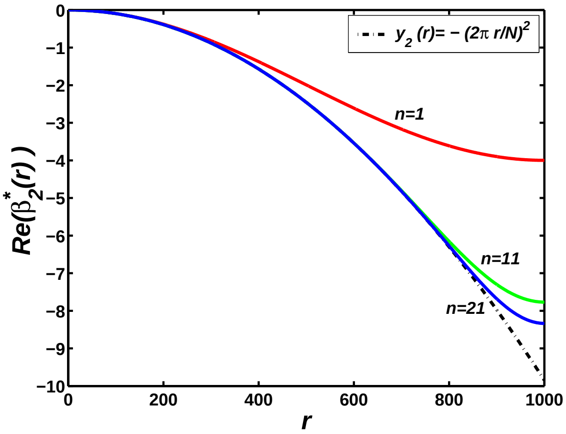


Figure 4: The spectra of various sequences $\alpha_m^{(2)}(n)$ at $N = 2000$.

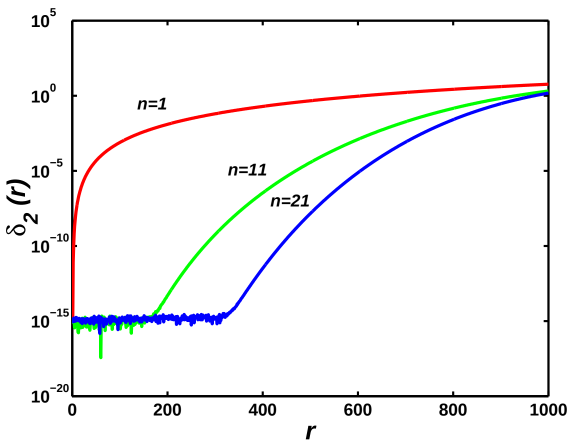


Figure 5: The function $\delta_2(r) = \mathfrak{Re}\big(\beta_2^*(r)\big) - y_2(r)$ versus r for different n.

developed method we have calculated the first derivative of the function $y(x) = \sin x$. The obtained result has been studied in the limiting case $n \to \infty$. We have examined the spectral characteristics of the weight coefficients sequence of the numerical differentiation formulae for the different number of the interpolation points. The performed analysis has allowed one to study the accuracy of the numerical differentiation carried out with the help of the developed method. It has been found that the derived formulae of numerical differentiation posses the high accuracy in a rather wide range of the spatial frequencies. As it has been shown in this paper, the formulae for the derivatives finding gave correct results in the case of large number of interpolation points. Thus, the developed method can be useful in lattice simulation of quantum fields [8]. To get the exact results at calculations on lattices one has to use nets with the big number of points. Derivatives which one encounters in theories of quantum fields, as a rule, do not exceed the second order. Therefore, the formulae obtained in this article could be of use in carrying out mentioned above research.

Acknowledgments

This research was supported by grant of Russian Science Support Foundation. The author is indebted to Sergey I. Dvornikov for helpful discussions.

A Algorithm for computation of weight coefficients $\alpha_m^{(1)}(n)$

In this appendix we present the algorithm for the computation of the coefficients $\alpha_m^{(1)}(n)$ on the MATLAB 6.5 programming language.

```
N = 2000;                       N is the size of the array
n = 11;                         n is the order of the differentiating sequence
α = zeros(1, N);                α is the array of the weight coefficients
k1 = 2; k2 = N;
for m = 1 : n
    r1 = 1;
    for k = 1 : n
        if k == m
            r2 = 1;
        else
            r2 = 1 − (m/k)^2;
        end
        r1 = r1 ∗ r2;
    end
    r1 = 1/(2 ∗ r1 ∗ m);
    α(1, k1) = −r1; k1 = k1 + 1;
    α(1, k2) = r1; k2 = k2 − 1;
end
```

References

[1] B. P. Demidovich and I. A. Maron, *Foundations of Computational Mathematics* (2nd ed.), Fiz. Mat. Lit., Moscow, 1963.

[2] N. N. Kalitkin, *Numerical Methods*, Nauka, Moscow, 1978.

[3] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, National Bureau of Standards, Washington D. C., 1964.

[4] W. G. Bickley, Formulae for numerical differentiation, *Math. Gaz.*, 25, 19–27, (1941).

[5] B. Fornberg, Generation of Finite Difference Formulas on Arbitrary Spaced Grids, *Math. Comp.*, 51(184), 699–706, (1988).

[6] V. I. Smirnov, *Course of higher mathematics, vol. 2* (16th ed.), Fiz. Mat. Lit., Moscow, 1958, p. 404.

[7] S. M. Kay and S. L. Marple, Jr., Spectrum Analysis – A Modern Perspective, in *Proceedings of the IEEE, vol. 69* (H. Freitag, ed.), IEEE Inc., 1981, pp. 1380–1419.

[8] M. Creutz, *Quarks, Gluons and Lattices*, Cambridge University Press, London, 1985.

Instructions to Contributors
Journal of Concrete and Applicable Mathematics

A quartely international publication of Eudoxus Press, LLC, of TN.

Editor in Chief: George Anastassiou
Department of Mathematical Sciences
University of Memphis
Memphis, TN 38152-3240, U.S.A.

1. Manuscripts hard copies in triplicate, and in English, should be submitted to the Editor-in-Chief:

Prof.George A. Anastassiou
Department of Mathematical Sciences
The University of Memphis
Memphis,TN 38152, USA.
Tel.001. 901.678.3144
e-mail:ganastss@memphis.edu.

Authors may want to recommend an associate editor the most related to the submission to possibly handle it.
Also authors may want to submit a list of six possible referees, to be used in case we cannot find related referees by ourselves.
2. Manuscripts should be typed using any of TEX,LaTEX,AMS-TEX,or AMS-LaTEX and according to EUDOXUS PRESS, LLC. LATEX STYLE FILE. (VISIT www.msci.memphis.edu/~ganastss/jcaam / to save a copy of the style file.)They should be carefully prepared in all respects. Submitted copies should be brightly printed (not dot-matrix), double spaced, in ten point type size, on one side high quality paper 8(1/2)x11 inch. Manuscripts should have generous margins on all sides and should not exceed 24 pages.
3. Submission is a representation that the manuscript has not been published previously in this or any other similar form and is not currently under consideration for publication elsewhere. A statement transferring from the authors(or their employers,if they hold the copyright) to Eudoxus Press, LLC, will be required before the manuscript can be accepted for publication.The Editor-in-Chief will supply the necessary forms for this transfer.Such a written transfer of copyright,which previously was assumed to be implicit in the act of submitting a manuscript,is necessary under the U.S.Copyright Law in order for the publisher to carry through the dissemination of research results and reviews as widely and effective as possible.
4. The paper starts with the title of the article, author's name(s) (no titles or degrees), author's affiliation(s) and e-mail addresses. The affiliation should comprise the department, institution (usually university or company), city, state (and/or nation) and mail code.
The following items, 5 and 6, should be on page no. 1 of the paper.

5. An abstract is to be provided, preferably no longer than 150 words.
6. A list of 5 key words is to be provided directly below the abstract. Key words should express the precise content of the manuscript, as they are used for indexing purposes. The main body of the paper should begin on page no. 1, if possible.
7. All sections should be numbered with Arabic numerals (such as: 1. INTRODUCTION) .
Subsections should be identified with section and subsection numbers (such as 6.1. Second-Value Subheading).
If applicable, an independent single-number system (one for each category) should be used to label all theorems, lemmas, propositions, corrolaries, definitions, remarks, examples, etc. The label (such as Lemma 7) should be typed with paragraph indentation, followed by a period and the lemma itself.
8. Mathematical notation must be typeset. Equations should be numbered consecutively with Arabic numerals in parentheses placed flush right,and should be thusly referred to in the text [such as Eqs.(2) and (5)]. The running title must be placed at the top of even numbered pages and the first author's name, et al., must be placed at the top of the odd numbed pages.
9. Illustrations (photographs, drawings, diagrams, and charts) are to be numbered in one consecutive series of Arabic numerals. The captions for illustrations should be typed double space. All illustrations, charts, tables, etc., must be embedded in the body of the manuscript in proper, final, print position. In particular, manuscript, source, and PDF file version must be at camera ready stage for publication or they cannot be considered.
Tables are to be numbered (with Roman numerals) and referred to by number in the text. Center the title above the table, and type explanatory footnotes (indicated by superscript lowercase letters) below the table.
10. List references alphabetically at the end of the paper and number them consecutively. Each must be cited in the text by the appropriate Arabic numeral in square brackets on the baseline.

References should include (in the following order):
initials of first and middle name, last name of author(s)
title of article,
name of publication, volume number, inclusive pages, and year of publication.

Authors should follow these examples:

Journal Article

1. H.H.Gonska,Degree of simultaneous approximation of bivariate functions by Gordon operators, (journal name in italics) *J. Approx. Theory*, 62,170-191(1990).

Book

2. G.G.Lorentz, (title of book in italics) *Bernstein Polynomials* (2nd ed.), Chelsea,New York,1986.

Contribution to a Book

3. M.K.Khan, Approximation properties of beta operators,in(title of book in italics) *Progress in Approximation Theory* (P.Nevai and A.Pinkus,eds.), Academic Press, New York,1991,pp.483-495.

11. All acknowledgements (including those for a grant and financial support) should occur in one paragraph that directly precedes the References section.
12. Footnotes should be avoided. When their use is absolutely necessary, footnotes should be numbered consecutively using Arabic numerals and should be typed at the bottom of the page to which they refer. Place a line above the footnote, so that it is set

off from the text. Use the appropriate superscript numeral for citation in the text.
13. After each revision is made please again submit three hard copies of the revised manuscript, including in the final one. And after a manuscript has been accepted for publication and with all revisions incorporated, manuscripts, including the TEX/LaTex source file and the PDF file, are to be submitted to the Editor's Office on a personal-computer disk, 3.5 inch size. Label the disk with clearly written identifying information and properly ship, such as:

Your name, title of article, kind of computer used, kind of software and version number, disk format and files names of article, as well as abbreviated journal name.

Package the disk in a disk mailer or protective cardboard. Make sure contents of disks are identical with the ones of final hard copies submitted!

Note: The Editor's Office cannot accept the disk without the accompanying matching hard copies of manuscript. No e-mail final submissions are allowed! The disk submission must be used.

14. Effective 1 Nov. 2005 the journal's page charges are $10.00 per PDF file page. Upon acceptance of the paper an invoice will be sent to the contact author. The fee payment will be due one month from the invoice date. The article will proceed to publication only after the fee is paid. The charges are to be sent, by money order or certified check, in US dollars, payable to Eudoxus Press, LLC, to the address shown in the Scope and
Prices section.

No galleys will be sent and the contact author will receive one(1) electronic copy of the journal issue in which the article appears.

15. This journal will consider for publication only papers that contain proofs for their listed results.

TABLE OF CONTENTS,JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS,VOL.5,NO.1,2007

VOLUME 5,NUMBER 2 APRIL 2007

ISSN:1548-5390 PRINT,1559-176X ONLINE

JOURNAL

OF CONCRETE

AND APPLICABLE

MATHEMATICS

EUDOXUS PRESS,LLC

SCOPE AND PRICES OF THE JOURNAL
Journal of Concrete and Applicable Mathematics

A quartely international publication of **Eudoxus Press,LLC**

Editor in Chief: George Anastassiou
Department of Mathematical Sciences,
University of Memphis
Memphis, TN 38152, U.S.A.
ganastss@memphis.edu

The main purpose of the "Journal of Concrete and Applicable Mathematics" is to publish high quality original research articles from all subareas of Non-Pure and/or Applicable Mathematics and its many real life applications, as well connections to other areas of Mathematical Sciences, as long as they are presented in a Concrete way. It welcomes also related research survey articles and book reviews.A sample list of connected mathematical areas with this publication includes and is not restricted to: Applied Analysis, Applied Functional Analysis, Probability theory, Stochastic Processes, Approximation Theory, O.D.E, P.D.E, Wavelet, Neural Networks,Difference Equations, Summability, Fractals, Special Functions, Splines, Asymptotic Analysis, Fractional Analysis, Inequalities, Moment Theory, Numerical Functional Analysis,Tomography, Asymptotic Expansions, Fourier Analysis, Applied Harmonic Analysis, Integral Equations, Signal Analysis, Numerical Analysis, Optimization, Operations Research, Linear Programming, Fuzzyness, Mathematical Finance, Stochastic Analysis, Game Theory, Math.Physics aspects, Applied Real and Complex Analysis, Computational Number Theory, Graph Theory, Combinatorics,
Computer Science Math.related topics,combinations of the above, etc. In general any kind of Concretely presented Mathematics which is Applicable fits to the scope of this journal.
Working Concretely and in Applicable Mathematics has become a main trend in many recent years,so we can understand better and deeper and solve the important problems of our real and scientific world.
"Journal of Concrete and Applicable Mathematics" is a peer- reviewed International Quarterly Journal.
We are calling for papers for possible publication. The contributor should send three copies of the contribution to the editor in-Chief typed in TEX, LATEX double spaced. [See: Instructions to Contributors]

Journal of Concrete and Applicable Mathematics(JCAAM)
ISSN:1548-5390 PRINT, 1559-176X ONLINE.
is published in January,April,July and October of each year by
EUDOXUS PRESS,LLC,
1424 Beaver Trail Drive,Cordova,TN38016,USA,
Tel.001-901-751-3553
anastassioug@yahoo.com
http://www.EudoxusPress.com.
Visit also www.msci.memphis.edu/~ganastss/jcaam.
Webmaster:Ray Clapsadle

Annual Subscription Current Prices:For USA and Canada,Institutional:Print $250,Electronic $220,Print and Electronic $310.Individual:Print $77,Electronic $60,Print

&Electronic $110.For any other part of the world add $25 more to the above prices for Print.
Single article PDF file for individual $8.Single issue in PDF form for individual $25.
The journal carries page charges $10 per page of the pdf file of an article,payable upon acceptance of the article within one month and before publication.
No credit card payments.Only certified check,money order or international check in US dollars are acceptable.
Combination orders of any two from JoCAAA,JCAAM,JAFA receive 25% discount,all three receive 30% discount.

JCAAM is reviewed and abstracted by AMS Mathematical Reviews,MATHSCI,and Zentralblaat MATH.

JCAAM IS A JOURNAL OF RAPID PUBLICATION

9) A.M.Fink
Department of Mathematics
Iowa State University
Ames,IA 50011-0001,USA
tel.515-294-8150
fink@math.iastate.edu
Inequalities,Ordinary Differential Equations

10) Sorin Gal
Department of Mathematics
University of Oradea
Str.Armatei Romane 5
3700 Oradea,Romania
galso@uoradea.ro
Approximation Th.,Fuzzyness,Complex Analysis

11) Jerome A.Goldstein
Department of Mathematical Sciences
The University of Memphis,
Memphis,TN 38152,USA
tel.901-678-2484
jgoldste@memphis.edu
Partial Differential Equations, Semigroups of Operators

12) Heiner H.Gonska
Department of Mathematics
University of Duisburg
Duisburg,D-47048
Germany
tel.0049-203-379-3542 office
gonska@informatik.uni-duisburg.de
Approximation Th.,Computer Aided Geometric Design

13) Dmitry Khavinson
Department of Mathematical Sciences
University of Arkansas
Fayetteville,AR 72701,USA
tel.(479)575-6331,fax(479)575-8630
dmitry@uark.edu
Potential Th.,Complex Analysis,Holomorphic PDE,
Approximation Th.,Function Th.

14) Virginia S.Kiryakova
Institute of Mathematics and Informatics
Bulgarian Academy of Sciences
Sofia 1090,Bulgaria
virginia@diogenes.bg
Special Functions,Integral Transforms, Fractional Calculus

15) Hans-Bernd Knoop

29) Bl. Sendov
Institute of Mathematics and Informatics
Bulgarian Academy of Sciences
Sofia 1090,Bulgaria
bsendov@bas.bg
Approximation Th.,Geometry of Polynomials, Image Compression

30) Igor Shevchuk
Faculty of Mathematics and Mechanics
National Taras Shevchenko
University of Kyiv
252017 Kyiv
UKRAINE
shevchuk@univ.kiev.ua
Approximation Theory

31) H.M.Srivastava
Department of Mathematics and Statistics
University of Victoria
Victoria,British Columbia V8W 3P4
Canada
tel.250-721-7455 office,250-477-6960 home,
fax 250-721-8962
harimsri@math.uvic.ca
Real and Complex Analysis,Fractional Calculus and Appl.,
Integral Equations and Transforms,Higher Transcendental
Functions and Appl.,q-Series and q-Polynomials,
Analytic Number Th.

32) Ferenc Szidarovszky
Dept.Systems and Industrial Engineering
The University of Arizona
Engineering Building,111
PO.Box 210020
Tucson,AZ 85721-0020,USA
szidar@sie.arizona.edu
Numerical Methods,Game Th.,Dynamic Systems,
Multicriteria Decision making,
Conflict Resolution,Applications
in Economics and Natural Resources
Management

33) Gancho Tachev
Dept.of Mathematics
Univ.of Architecture,Civil Eng. and Geodesy
1 Hr.Smirnenski blvd
BG-1421 Sofia,Bulgaria
gtt_fte@uacg.bg
Approximation Theory

34) Manfred Tasche
Department of Mathematics

Institute of Mathematics
Gerhard Mercator University
D-47048 Duisburg
Germany
tel.0049-203-379-2676
knoop@math.uni-duisburg.de
Approximation Theory,Interpolation

16) Jerry Koliha
Dept. of Mathematics & Statistics
University of Melbourne
VIC 3010,Melbourne
Australia
koliha@unimelb.edu.au
Inequalities,Operator Theory,
Matrix Analysis,Generalized Inverses

17) Mustafa Kulenovic
Department of Mathematics
University of Rhode Island
Kingston,RI 02881,USA
kulenm@math.uri.edu
Differential and Difference Equations

18) Gerassimos Ladas
Department of Mathematics
University of Rhode Island
Kingston,RI 02881,USA
gladas@math.uri.edu
Differential and Difference Equations

19) Rupert Lasser
Institut fur Biomathematik & Biomertie,GSF
-National Research Center for environment
and health
Ingolstaedter landstr.1
D-85764 Neuherberg,Germany
lasser@gsf.de
Orthogonal Polynomials,Fourier Analysis,
Mathematical Biology

University of Rostock
D-18051 Rostock
Germany
manfred.tasche@mathematik.uni-rostock.de
Approximation Th.,Wavelet,Fourier Analysis,
Numerical Methods,Signal Processing,
Image Processing,Harmonic Analysis

35) Chris P.Tsokos
Department of Mathematics
University of South Florida
4202 E.Fowler Ave.,PHY 114
Tampa,FL 33620-5700,USA
profcpt@math.usf.edu,profcpt@chuma1.cas.usf.edu
Stochastic Systems,Biomathematics,
Environmental Systems,Reliability Th.

36) Lutz Volkmann
Lehrstuhl II fuer Mathematik
RWTH-Aachen
Templergraben 55
D-52062 Aachen
Germany
volkm@math2.rwth-aachen.de
Complex Analysis,Combinatorics,Graph Theory

JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS,VOL.5,NO.2,101-113,2007,COPYRIGHT 2007 EUDOXUS PRESS,LLC

On the Difference Equation $x_{n+1} = \dfrac{\alpha x_{n-k}}{\beta + \gamma \prod_{i=0}^{k} x_{n-i}}$

E.M. Elabbasy[1], H. El-Metwally[2], and E.M. Elsayed[3].
Department of Mathematics, Faculty of Science,
Mansoura University, Mansoura 35516, Egypt.
[1]E-mail: emelabbasy@mans.edu.eg.
[2]E-mail: helmetwally@mans.edu.eg.
[3]E-mail: emelsayed@mans.edu.eg.

February 25, 2006

Abstract

In this paper we investigate some qualitative behavior of the solutions of the recursive sequence

$$x_{n+1} = \frac{\alpha x_{n-k}}{\beta + \gamma \prod_{i=0}^{k} x_{n-i}}, \quad n = 0, 1, ...$$

where the parameters α, β and γ are non-negative real numbers and the initial values $x_{-k},\ x_{-k+1}, ...,\ x_0$ are arbitrary positive numbers.

Keywords: difference equations, stability, periodic solutions.
Mathematics Subject Classification: 39A10

1 Introduction

In this paper we deal with some properties of the solutions of the recursive sequence

$$x_{n+1} = \frac{\alpha x_{n-k}}{\beta + \gamma \prod_{i=0}^{k} x_{n-i}}, \quad n = 0, 1, ... \tag{1}$$

where the parameters α, β and γ are positive real numbers and the initial conditions $x_{-k},\ x_{-k+1}, ...,\ x_0$ are arbitrary positive numbers.

Recently there has been a lot of interest in studying the global attractivity, boundedness character and the periodic nature of nonlinear difference equations. For some results in this area, see for example [4-6].

Cinar [1-3] has investigated the solutions of some special cases of Eq.(1) when $k = 1$.

The study of the nonlinear rational difference equations of a higher order is quite challenging and rewarding, and the results about these equations offer prototypes towards the development of the basic theory of the global behavior of nonlinear difference equations of a big order.

Throughout this work we set $A = x_{-k}x_{-k+1}...x_{-1}x_0$.

Here, we recall some notations and results which will be useful in our investigation of Eq.(1).

Let I be some interval of real numbers and let f be a continuous function defined on I^{k+1}. Then, for initial conditions $x_{-k}, x_{-k+1}, ..., x_0 \in I$, it is easy to see that the difference equation

$$x_{n+1} = f(x_n, x_{n-1}, ..., x_{n-k}), \quad n = 0, 1, ..., \tag{2}$$

has a unique solution $\{x_n\}_{n=-k}^{\infty}$.

A point $\overline{x} \in I$ is called an equilibrium point of Eq.(2) if

$$\overline{x} = f(\overline{x}, \overline{x}, ..., \overline{x}).$$

That is, $x_n = \overline{x}$ for $n \geq 0$, is a solution of Eq.(2), or equivalently, $\overline{x}$ is a fixed point of f.

Definition 1 *(Stability)*

(i) The equilibrium point $\overline{x}$ of Eq.(2) is locally stable if for every $\epsilon > 0$, there exists $\delta > 0$ such that for all $x_{-k}, x_{-k+1}, ..., x_{-1}, x_0 \in I$ with

$$|x_{-k} - \overline{x}| + |x_{-k+1} - \overline{x}| + ... + |x_0 - \overline{x}| < \delta,$$

we have

$$|x_n - \overline{x}| < \epsilon \qquad \text{for all} \quad n \geq -k.$$

(ii) The equilibrium point $\overline{x}$ of Eq.(2) is locally asymptotically stable if $\overline{x}$ is locally stable solution of Eq.(2) and there exists $\gamma > 0$, such that for all $x_{-k}, x_{-k+1}, ...,$ $x_{-1}, x_0 \in I$ with

$$|x_{-k} - \overline{x}| + |x_{-k+1} - \overline{x}| + ... + |x_0 - \overline{x}| < \gamma,$$

we have

$$\lim_{n\to\infty} x_n = \overline{x}.$$

(iii) The equilibrium point $\overline{x}$ of Eq.(2) is global attractor if for all $x_{-k}, x_{-k+1}, ...,$ $x_{-1}, x_0 \in I$, we have

$$\lim_{n\to\infty} x_n = \overline{x}.$$

(iv) The equilibrium point $\overline{x}$ of Eq.(2) is globally asymptotically stable if $\overline{x}$ is locally stable, and $\overline{x}$ is also a global attractor of Eq.(2).

(v) The equilibrium point $\overline{x}$ of Eq.(2) is unstable if $\overline{x}$ is not locally stable.

The linearized equation of Eq.(2) about the equilibrium $\overline{x}$ is the linear difference equation

$$y_{n+1} = \sum_{i=0}^{k} \frac{\partial F(\overline{x}, \overline{x}, ..., \overline{x})}{\partial x_{n-i}} y_{n-i}, \quad n = 0, 1, ...$$

Its characteristic equation is

$$\lambda^{k+1} - \sum_{i=0}^{k} \frac{\partial f(\overline{x}, \overline{x}, ..., \overline{x})}{\partial x_{n-i}} \lambda^{k-i} = 0$$

Theorem A [6]: Assume that $p_i \in R$, $i = 1, 2, ...$ and $k \in \{0, 1, 2, ...\}$. Then

$$\sum_{i=1}^{k} |p_i| < 1$$

is a sufficient condition for the asymptotic stability of the difference equation

$$x_{n+k} + p_1 x_{n+k-1} + ... + p_k x_n = 0, \quad n = 0, 1, ...$$

2 Local Stability of the Equilibrium Points

In this section we study the local stability character of the solutions of Eq.(1).

The equilibrium points of Eq.(1) are given by the relation

$$\overline{x} = \frac{\alpha \overline{x}}{\beta + \gamma \overline{x}^{k+1}},$$

which gives

$$\overline{x} = 0 \quad or \quad \overline{x} = \left[\frac{\alpha - \beta}{\gamma}\right]^{\frac{1}{k+1}}.$$

Note that if $\alpha > \beta$, then Eq.(1) has a positive equilibrium point.

Let $f : (0, \infty)^{k+1} \longrightarrow (0, \infty)$ be a function defined by

$$f(x_n, x_{n-1}, ..., x_{n-k}) = \frac{\alpha x_{n-k}}{\beta + \gamma \prod_{i=0}^{k} x_{n-i}}. \tag{3}$$

Theorem 1 *The following statements are true*

(1) If $\alpha < \beta$, then the equilibrium point $\overline{x} = 0$ of Eq.(1) is locally stable.

(2) If $\alpha > \beta$, then the positive equilibrium point $\overline{x} = \left[\frac{\alpha-\beta}{\gamma}\right]^{\frac{1}{k+1}}$ of Eq.(1) is locally stable when $k \neq 1$.

Proof: (1) If $\alpha < \beta$, then we see from Eq.(3) that

$$\begin{aligned} \frac{\partial f(0,0,...,0)}{\partial x_{n-i}} &= 0, \quad i \neq k, \\ \frac{\partial f(0,0,...,0)}{\partial x_{n-k}} &= \frac{\alpha}{\beta}. \end{aligned}$$

Then the linearized equation associated with Eq.(1) about 0 is

$$y_{n+1} - \frac{\alpha}{\beta} y_{n-k} = 0 \tag{4}$$

whose characteristic equation is

$$\lambda^{k+1} - \frac{\alpha}{\beta} = 0. \tag{5}$$

It follows by Theorem A that, Eq.(4) is asymptotically stable. Then the equilibrium point $\overline{x} = 0$ of Eq.(1) is locally stable.

(2) If $\alpha > \beta$, then we see from Eq.(3) that

$$\begin{aligned} \frac{\partial f(\overline{x},\overline{x},...,\overline{x})}{\partial x_{n-i}} &= -\frac{\alpha-\beta}{\alpha}, \quad i \neq k, \\ \frac{\partial f(\overline{x},\overline{x},...,\overline{x})}{\partial x_{n-k}} &= \frac{\beta}{\alpha}. \end{aligned}$$

Then the linearized equation of Eq.(1) about $\overline{x}$ is

$$y_{n+1} + \frac{\alpha-\beta}{\alpha} y_n + \frac{\alpha-\beta}{\alpha} y_{n-1} + ... + \frac{\alpha-\beta}{\alpha} y_{n-k+1} - \frac{\beta}{\alpha} y_{n-k} = 0 \tag{6}$$

whose characteristic equation is

$$\lambda^{k+1} + \frac{\alpha-\beta}{\alpha}\lambda^k + \frac{\alpha-\beta}{\alpha}\lambda^{k-1} + ... + \frac{\alpha-\beta}{\alpha}\lambda - \frac{\beta}{\alpha} = 0. \tag{7}$$

It follows by Theorem A that, Eq.(6) is asymptotically stable if all roots of Eq.(7) lie in the open disc $|\lambda| < 1$ that is if

$$\left|\frac{\alpha-\beta}{\alpha}\right| + \left|\frac{\alpha-\beta}{\alpha}\right| + ... + \left|\frac{\alpha-\beta}{\alpha}\right| + \left|\frac{\beta}{\alpha}\right| < 1$$

which is true if

$$k\left[\frac{\alpha-\beta}{\alpha}\right] + \frac{\beta}{\alpha} < 1 \;\Leftrightarrow\; \frac{\beta}{\alpha}(1-k) < (1-k)$$

which is always satisfied where $\beta < \alpha$ and where $k \neq 1$. The proof is complete.

3 Existence of Period (k+1) Solutions

In this section we study the existence of period $(k+1)$ solutions of Eq.(1).
Note that if we choose all the initial values $\{x_{-k}, x_{-k+1}, x_{-k+2}, ..., x_{-1}, x_0\}$ equal to zero, then Eq.(1) has only the trivial solution.

Theorem 2 *Eq.(1) has positive prime period $(k+1)$ solutions if and only if*

$$(\beta + \gamma A - \alpha) = 0, \tag{8}$$

where $A = \prod_{i=0}^{k} x_{-i}$.

Proof: First suppose that there exists a prime period $(k+1)$ solution of Eq.(1) of the form

$$..., x_{-k}, x_{-k+1}, x_{-k+2}, ..., x_{-1}, x_0, x_{-k}, x_{-k+1}, x_{-k+2}, ..., x_{-1}, x_0, ...$$

That is $x_{N+1} = x_{N-k}$ for $N \geq 0$. We will prove that (8) holds.
From Eq.(1), we see that

$$\begin{aligned} x_{-k} &= x_1 = \frac{\alpha x_{-k}}{\beta + \gamma A}, \quad x_{-k+1} = x_2 = \frac{\alpha x_{-k+1}}{\beta + \gamma A}, \quad x_{-k+2} = x_3 = \frac{\alpha x_{-k+2}}{\beta + \gamma A}, ..., \\ x_{-1} &= x_k = \frac{\alpha x_{-1}}{\beta + \gamma A}, \quad x_0 = x_{k+1} = \frac{\alpha x_0}{\beta + \gamma A}. \end{aligned}$$

Since it has to be there is at least one of the initial values $\{x_{-k}, x_{-k+1}, ..., x_0\}$ does not equal zero, then Condition (8) is satisfied.
Second suppose that (8) is true. We will show that Eq.(1) has a prime period $(k+1)$ solution.

It follows from Eq.(1) that

$$\begin{aligned} x_1 &= \frac{\alpha x_{-k}}{\beta + \gamma A} = x_{-k}, \quad x_2 = \frac{\alpha x_{-k+1}}{\beta + \gamma A} = x_{-k+1}, \quad x_3 = \frac{\alpha x_{-k+2}}{\beta + \gamma A} = x_{-k+2}, ..., \\ x_k &= \frac{\alpha x_{-1}}{\beta + \gamma A} = x_{-1}, \quad x_{k+1} = \frac{\alpha x_0}{\beta + \gamma A} = x_0, \end{aligned}$$

The proof is complete.

4 The Solution Form of Eq.(1)

In this section we give a specific form of the solutions of Eq.(1)

Note that the change of variables $x_{n+1} = \left(\frac{\beta}{\gamma}\right)^{\frac{1}{(k+1)}} y_{n+1}$ reduces Eq.(1) to

$$y_{n+1} = \frac{a y_{n-k}}{1 + \prod_{i=0}^{k} y_{n-i}}, \quad n = 0, 1, ... , \tag{9}$$

where $a = \dfrac{\alpha}{\beta}$.

Theorem 3 *Let $\{y_n\}_{n=-k}^{\infty}$ be a solution of Eq.(9). Then for $n = 0, 1, ...$*

$$y_{(k+1)n+s} = \frac{a^{n+1} y_{s-(k+1)} \Pi_{i=0}^{n} \left(1 + \Sigma_{j=0}^{(k+1)i+(s-2)} Ba^j\right)}{\Pi_{i=0}^{n} \left(1 + \Sigma_{j=0}^{(k+1)i+(s-1)} Ba^j\right)}, \ s = 1, 2, ..., k+1,$$

where $B = y_{-k} y_{-k+1} ... y_{-1} y_0$ and $\Pi_{i=1}^{0} \left(1 + \Sigma_{j=0}^{(k+1)i-1} Ba^j\right) = 1$.

Proof: For $n = 0$ the result holds. Now suppose that $n > 0$ and that our assumption holds for $n - 1$. That is;

$$\begin{aligned}
y_{(k+1)n-k} &= \frac{a^n y_{-k} \Pi_{i=1}^{n-1} \left(1 + \Sigma_{j=0}^{(k+1)i-1} Ba^j\right)}{\Pi_{i=0}^{n-1} \left(1 + \Sigma_{j=0}^{(k+1)i} Ba^j\right)}, \\
y_{(k+1)n-k+1} &= \frac{a^n y_{-k+1} \Pi_{i=0}^{n-1} \left(1 + \Sigma_{j=0}^{(k+1)i} Ba^j\right)}{\Pi_{i=0}^{n-1} \left(1 + \Sigma_{j=0}^{(k+1)i+1} Ba^j\right)}, \\
y_{(k+1)n-k+2} &= \frac{a^n y_{-k+2} \Pi_{i=0}^{n-1} \left(1 + \Sigma_{j=0}^{(k+1)i+1} Ba^j\right)}{\Pi_{i=0}^{n-1} \left(1 + \Sigma_{j=0}^{(k+1)i+2} Ba^j\right)}, ..., \\
y_{(k+1)n-1} &= \frac{a^n y_{-1} \Pi_{i=0}^{n-1} \left(1 + \Sigma_{j=0}^{(k+1)i+k-2} Ba^j\right)}{\Pi_{i=0}^{n-1} \left(1 + \Sigma_{j=0}^{(k+1)i+k-1} Ba^j\right)},
\end{aligned}$$

and

$$y_{(k+1)n} = \frac{a^n y_0 \Pi_{i=0}^{n-1} \left(1 + \Sigma_{j=0}^{(k+1)i+k-1} Ba^j\right)}{\Pi_{i=0}^{n-1} \left(1 + \Sigma_{j=0}^{(k+1)i+k} Ba^j\right)}.$$

Now, it follows from Eq.(9) that

$$y_{(k+1)n+1} = \frac{a y_{(k+1)n-k}}{1 + \prod_{i=0}^{k} y_{(k+1)n-i}} = \frac{\dfrac{a^{n+1} y_{-k} \Pi_{i=1}^{n-1} \left(1 + \Sigma_{j=0}^{(k+1)i-1} Ba^j\right)}{\Pi_{i=0}^{n-1} \left(1 + \Sigma_{j=0}^{(k+1)i} Ba^j\right)}}{1 + B \dfrac{a^{(k+1)n} \Pi_{i=1}^{n-1} \left(1 + \Sigma_{j=0}^{(k+1)i-1} Ba^j\right)}{\Pi_{i=0}^{n-1} \left(1 + \Sigma_{j=0}^{(k+1)i+k} Ba^j\right)}}$$

$$= \frac{a^{n+1}y_{-k}\Pi_{i=1}^{n-1}\left(1+\Sigma_{j=0}^{(k+1)i-1}Ba^j\right)}{\Pi_{i=0}^{n-1}\left(1+\Sigma_{j=0}^{(k+1)i}Ba^j\right)\left\{1+B\frac{a^{(k+1)n}}{\left(1+\Sigma_{j=0}^{(k+1)(n-1)+k}Ba^j\right)}\right\}}$$

$$= \frac{a^{n+1}y_{-k}\Pi_{i=1}^{n-1}\left(1+\Sigma_{j=0}^{(k+1)i-1}Ba^j\right)\left(1+\Sigma_{j=0}^{(k+1)(n-1)+k}Ba^j\right)}{\Pi_{i=0}^{n-1}\left(1+\Sigma_{j=0}^{(k+1)i}Ba^j\right)\left\{\left(1+\Sigma_{j=0}^{(k+1)(n-1)+k}Ba^j\right)+Ba^{(k+1)n}\right\}}$$

$$= \frac{a^{n+1}y_{-k}\Pi_{i=1}^{n-1}\left(1+\Sigma_{j=0}^{(k+1)i-1}Ba^j\right)\left(1+\Sigma_{j=0}^{(k+1)n-1}Ba^j\right)}{\Pi_{i=0}^{n-1}\left(1+\Sigma_{j=0}^{(k+1)i}Ba^j\right)\left\{\left(1+\Sigma_{j=0}^{(k+1)n-1}Ba^j\right)+Ba^{(k+1)n}\right\}}$$

$$= \frac{a^{n+1}y_{-k}\Pi_{i=1}^{n}\left(1+\Sigma_{j=0}^{(k+1)i-1}Ba^j\right)}{\Pi_{i=0}^{n-1}\left(1+\Sigma_{j=0}^{(k+1)i}Ba^j\right)\left(1+\Sigma_{j=0}^{(k+1)n}Ba^j\right)}.$$

Hence, we have

$$y_{(k+1)n+1} = \frac{a^{n+1}y_{-k}\Pi_{i=1}^{n}\left(1+\Sigma_{j=0}^{(k+1)i-1}Ba^j\right)}{\Pi_{i=0}^{n}\left(1+\Sigma_{j=0}^{(k+1)i}Ba^j\right)}.$$

Similarly, by some simple computations, we get

$$y_{(k+1)n+s} = \frac{a^{n+1}y_{s-(k+1)}\Pi_{i=0}^{n}\left(1+\Sigma_{j=0}^{(k+1)i+(s-2)}Ba^j\right)}{\Pi_{i=0}^{n}\left(1+\Sigma_{j=0}^{(k+1)i+(s-1)}Ba^j\right)}, \quad s=2,3,...,k+1.$$

Hence, the proof is completed.

5 Boundedness of Solutions

Here we study the boundedness nature of the solutions of Eq.(1).

Theorem 4 *Every solution of Eq.(1) is bounded.*

Proof: Let $\{x_n\}_{n=-k}^{\infty}$ be a solution of Eq.(1), we consider the following two cases
(1) If $\alpha \leq \beta$. It follows from Eq.(1) that

$$x_{n+1} = \frac{\alpha x_{n-k}}{\beta+\gamma\prod_{i=0}^{k}x_{n-i}} \leq \frac{\alpha x_{n-k}}{\beta} \leq x_{n-k}.$$

Then the subsequences $\{x_{(k+1)n-k}\}_{n=0}^{\infty}$, $\{x_{(k+1)n-k+1}\}_{n=0}^{\infty}$, ..., $\{x_{(k+1)n-1}\}_{n=0}^{\infty}$ and

$\{x_{(k+1)n}\}_{n=0}^{\infty}$ are decreasing and so are bounded from above by $M = \max\{x_{-k}, x_{-k+1}, x_{-k+2}, ..., x_{-1}, x_0\}$.

(2) If $\alpha > \beta$, For the sake of contradiction, suppose that there exists a subsequence $\{x_{(k+1)n-k}\}_{n=0}^{\infty}$ is not bounded from above. Then from Eq.(1) we see for sufficiently large n that

$$\infty = \lim_{n\to\infty} x_{(k+1)n+1} = \lim_{n\to\infty} \frac{\alpha x_{(k+1)n-k}}{\beta + \gamma \prod_{i=0}^{k} x_{(k+1)n-i}} = \lim_{n\to\infty} \frac{\alpha}{\dfrac{\beta}{x_{(k+1)n-k}} + \gamma \prod_{i=0}^{k-1} x_{(k+1)n-i}}$$

$$= \lim_{n\to\infty} \frac{\alpha}{\gamma \prod_{i=0}^{k-1} x_{(k+1)n-i}} \tag{10}$$

It follows by Theorem 4 that the limit of the right hand side of (10) is bounded which is a contradiction, and so the proof of the theorem is complete.

6 Global Attractor of the Equilibrium Points of Eq.(1)

In this section we investigate the global attractivity character of solutions of Eq.(1). We give the following theorem which is a minor modification of Theorem A.0.7 in [8].

Theorem 5 *Let $[a, b]$ be an interval of real numbers and assume that*

$$f : [a, b]^{k+1} \to [a, b]$$

is a continuous function satisfying the following properties :

(a) $f(x_1, x_2, ..., x_{k+1})$ is non-increasing in the first (k) terms for each x_{k+1} in $[a, b]$ and non-decreasing in the last term for each x_i in $[a, b]$ for all $i = 1, 2, ..., k$;

(b) If $(m, M) \in [a, b] \times [a, b]$ is a solution of the system

$$M = f(m, m, ..., m, M) \quad and \quad m = f(M, M, ..., M, m)$$

implies

$$m = M.$$

Then Eq.(2) has a unique equilibrium $\overline{x} \in [a, b]$ and every solution of Eq.(2) converges to $\overline{x}$.

Proof: Set

$$m_0 = a \quad and \quad M_0 = b$$

and for each $i = 1, 2, ...$ set

$$M_i = f(m_{i-1}, m_{i-1}, ..., m_{i-1}, M_{i-1})$$

and

$$m_{i-1} = f(M_{i-1}, M_{i-1}, ..., M_{i-1}, m_{i-1}).$$

Now observe that for each $i \geq 0$,

$$a = m_0 \leq m_1 \leq ... \leq m_i \leq ... \leq M_i \leq ... \leq M_1 \leq M_0 = b,$$

and

$$m_i \leq x_p \leq M_i \quad for \quad p \geq (k+1)i + 1.$$

Set

$$m = \lim_{i\to\infty} m_i \quad and \quad M = \lim_{i\to\infty} M_i.$$

Then

$$M \geq \lim_{i\to\infty} \sup x_i \geq \lim \inf_{i\to\infty} x_i \geq m$$

and by the continuity of f,

$$M = f(m, m, ..., m, M) \quad and \quad m = f(M, M, ..., M, m).$$

In view of (b),

$$m = M = \overline{x},$$

from which the result follows.

Theorem 6 *The following statements are true*

(1) If $\alpha \leq \beta$, then the equilibrium point $\overline{x} = 0$ of Eq.(1) is global attractor.

(2) If $\alpha > \beta$, then the positive equilibrium point $\overline{x} = \left[\frac{\alpha-\beta}{\gamma}\right]^{\frac{1}{k+1}}$ of Eq.(1) is global attractor when $k \neq 1$.

Proof: (1) If $\alpha \leq \beta$, then the proof follows by Theorem 4.

(2) If $\alpha > \beta$, then we can easily see that the function $f(x_n, x_{n-1}, ..., x_{n-k})$ defined by Eq.(3) increasing in x_{n-k} and decreasing in the rest of arguments.

Suppose that (m, M) is a solution of the system

$$M = f(m, m, ..., m, M) \quad and \quad m = f(M, M, ..., M, m).$$

Then from Eq.(3), we see that

$$M = \frac{\alpha M}{\beta + \gamma M m^k}, \quad m = \frac{\alpha m}{\beta + \gamma M^k m}, \quad \Rightarrow \quad \beta + \gamma M m^k = \alpha, \quad \beta + \gamma M^k m = \alpha$$

$$M m^k = M^k m$$

Thus

$$M = m.$$

It follows by Theorem 6 that $\overline{x}$ is a global attractor of Eq.(1) and then the proof is complete.

7 Applications

7.1 On the Difference Equation $x_{n+1} = \dfrac{1}{\prod_{i=0}^{k} x_{n-i}}$

In this section we investigate the solutions of the recursive sequence

$$x_{n+1} = \frac{1}{\prod_{i=0}^{k} x_{n-i}}, \quad n = 0, 1, ... \tag{11}$$

where the initial conditions $x_{-k}, x_{-k+1}, ..., x_0$ are arbitrary real numbers.

Remark 1 *It is well known that every solution of the equation*

$$x_{n+1} = \frac{1}{x_{n-r}}, \qquad n = 0, 1, ...$$

is periodic with period $(2r + 2)$. The following theorem show that every solution of Eq.(11) is periodic with period equal to the average of the periods of the equations

$$x_{n+1} = \frac{1}{x_{n-j}}, \quad j = 0, 1, ..., k.$$

Theorem 7 *Let $\{x_n\}_{n=-k}^{\infty}$ be a solution of Eq.(11). Then every solution of Eq.(11) is periodic with period $(k+2)$. Moreover $\{x_n\}_{n=-k}^{\infty}$ be in the form*

$$\left\{..., \tfrac{1}{A}, x_{-k}, x_{-k+1}, ..., x_{-1}, x_0, \tfrac{1}{A}, x_{-k}, x_{-k+1}, ..., x_{-1}, x_0, ...\right\}.$$

Proof: From Eq.(11), we see that

$$\begin{aligned}
x_1 &= \frac{1}{\prod_{i=0}^{k} x_{-i}} = \frac{1}{A}, \quad x_2 = \frac{1}{\prod_{i=0}^{k} x_{1-i}} = \frac{A}{\prod_{i=1}^{k} x_{1-i}} = \frac{x_{-k}A}{x_{-k}\prod_{i=1}^{k} x_{1-i}} = x_{-k}, \\
x_3 &= \frac{1}{\prod_{i=0}^{k} x_{2-i}} = \frac{A}{x_{-k}\prod_{i=2}^{k} x_{2-i}} = x_{-k+1}, ..., x_{k-1} = \frac{1}{\prod_{i=0}^{k} x_{k-2-i}} = x_{-3}, \\
x_k &= \frac{1}{\prod_{i=0}^{k} x_{k-1-i}} = x_{-2}, \quad x_{k+1} = \frac{1}{\prod_{i=0}^{k} x_{k-i}} = x_{-1}, \quad x_{k+2} = \frac{1}{\prod_{i=0}^{k} x_{k+1-i}} = x_0,
\end{aligned}$$

Again it is easy to see that

$$\begin{aligned}
x_{k+3} &= \frac{1}{\prod_{i=0}^{k} x_{k+2-i}} = x_1 = \frac{1}{A}, \quad x_{k+4} = \frac{1}{\prod_{i=0}^{k} x_{k+3-i}} = x_{-2} = x_{-k}, ..., \\
x_{2k+4} &= \frac{1}{\prod_{i=0}^{k} x_{2k+3-i}} = x_{k+2} = x_0.
\end{aligned}$$

Hence, the proof is completed.

7.2 On the Difference Equation $x_{n+1} = \dfrac{x_{n-k}}{-1 + a\prod_{i=0}^{k} x_{n-i}}$

In this section we study the solutions of the recursive sequence

$$x_{n+1} = \frac{x_{n-k}}{-1 + a\prod_{i=0}^{k} x_{n-i}} \tag{12}$$

where the parameter a is real number and the initial conditions $x_{-k},\ x_{-k+1}, ...,\ x_0$ are arbitrary real numbers.
We study the following two cases of Eq.(12)
(I) When k is even. In this case we give the following theorem.

Theorem 8 *Let $\{x_n\}_{n=-k}^{\infty}$ be a solution of Eq.(12). Assume that $Aa \neq 1$. Then every solution of Eq.(12) is periodic with period $(2k+2)$. Moreover $\{x_n\}_{n=-k}^{\infty}$ takes the form*

$$\{..., \alpha_0, \alpha_1, ..., \alpha_k, x_{-k}, x_{-k+1}, ..., x_{-1}, x_0, \alpha_0, \alpha_1, ..., \alpha_k, x_{-k}, x_{-k+1}, ..., x_{-1}, x_0, ...\}.$$

where $\alpha_{2i} = \frac{x_{-k+2i}}{Aa-1}$, $\quad i = 0, 1, ...\frac{k}{2}$ *and* $\alpha_{2j+1} = (Aa-1)$, $\quad j = 0, 1, ...\frac{k}{2} - 1$.

Proof: From Eq.(12), we see that

$$\begin{aligned}
x_1 &= \frac{x_{-k}}{-1 + a\prod_{i=0}^{k} x_{-i}} = \frac{x_{-k}}{(Aa-1)}, \\
x_2 &= \frac{x_{-k+1}}{-1 + a\prod_{i=0}^{k} x_{1-i}} = \frac{x_{-k+1}}{-1 + \frac{Aa}{(Aa-1)}} = x_{-k+1}(Aa-1), \\
x_3 &= \frac{x_{-k+2}}{-1 + a\prod_{i=0}^{k} x_{2-i}} = \frac{x_{-k+2}}{-1 + Aa}, ..., x_k = \frac{x_{-1}}{-1 + a\prod_{i=0}^{k} x_{k-1-i}} = x_{-1}(Aa-1),
\end{aligned}$$

$$x_{k+1} = \frac{x_0}{-1 + a\prod_{i=0}^{k} x_{k-i}} = \frac{x_0}{-1 + Aa}.$$

Similarly to the above it is easy to show that

$$x_{k+2} = x_{-k},\ x_{k+3} = x_{-k+1}, ..., x_{2k+2} = x_{k+1},$$

By some other simple computations, it follows from Eq.(12) once again that

$$x_{2k+3} = x_1 = \frac{x_{-k}}{(Aa-1)},\ \ x_{2k+4} = x_2 = (Aa-1)x_{-k+1}, ..., x_{4k+4} = x_{2k+2} = x_0.$$

Hence, the proof is completed.

Remark 2 *If $Aa = 2$, then every solution of Eq.(12) is periodic with prime period $(k+1)$ and $\{x_n\}_{n=-k}^{\infty}$ takes the form*

$$\{..., x_{-k}, x_{-k+1}, ..., x_{-1}, x_0, x_{-k}, x_{-k+1}, ..., x_{-1}, x_0, ...\}.$$

(II) When k is odd
Now we study Eq.(12) whenever k is an odd integer number.

Theorem 9 *Assume that $Aa \neq 1$. Then Eq. (12) has unbounded solutions moreover for $n = 0, 1, ...$*

$$x_{(k+1)n+s} = \begin{cases} \frac{x_{-k-1+s}}{(-1+Aa)^{n+1}}, & s-odd \\ x_{-k-1+s}(-1+Aa)^{n+1}, & s-even \end{cases}, \quad s = 1, 2, ..., (k+1).$$

Proof: For $n = 0$ the result holds. Now suppose that $n > 0$ and that our assumption holds for $n-1$. We shall show that the result holds for n. From our assumption for $n-1$, we have

$$\begin{aligned} x_{(k+1)n-k} &= \frac{x_{-k}}{(-1+Aa)^n}, \quad x_{(k+1)n-k+1} = x_{-k+1}(-1+Aa)^n, \\ x_{(k+1)n-k+2} &= \frac{x_{-k+2}}{(-1+Aa)^n}, \quad x_{(k+1)n-k+3} = x_{-k+3}(-1+Aa)^n, ..., \\ x_{(k+1)n-2} &= x_{-2}(-1+Aa)^n, \ x_{(k+1)n-1} = \frac{x_{-1}}{(-1+Aa)^n}, \ x_{(k+1)n} = x_0(-1+Aa)^n, \end{aligned}$$

it follows from Eq.(12) that

$$x_{(k+1)n+1} = \frac{x_{-k}}{(-1+Aa)^n(-1+a\prod_{i=0}^{k} x_{(k+1)n-i})} = \frac{x_{-k}}{(-1+Aa)^{n+1}}$$

Sisimilarly

$$\begin{aligned} x_{(k+1)n+1} &= \frac{x_{-k}}{(-1+Aa)^{n+1}}, \quad x_{(k+1)n+2} = x_{-k+1}(-1+Aa)^{n+1}, \\ x_{(k+1)n+3} &= \frac{x_{-k+2}}{(-1+Aa)^{n+1}}, \quad x_{(k+1)n+4} = x_{-k+3}(-1+Aa)^{n+1}, ..., \\ x_{(k+1)n+(k-1)} &= x_{-2}(-1+Aa)^{n+1}, \ x_{(k+1)n+k} = \frac{x_{-1}}{(-1+Aa)^{n+1}}, \end{aligned}$$

and

$$x_{(k+1)n+(k+1)} = x_0(-1+A)^{n+1}.$$

Hence, the proof is completed.

References

[1] C. Cinar, On the positive solutions of the difference equation $x_{n+1} = \frac{x_{n-1}}{1+x_n x_{n-1}}$, Applied Mathematics and Computation, in press.

[2] C. Cinar, On the positive solutions of the difference equation $x_{n+1} = \frac{x_{n-1}}{1+ax_n x_{n-1}}$, Applied Mathematics and Computation, in press.

[3] C. Cinar, On the difference equation $x_{n+1} = \frac{x_{n-1}}{-1+x_n x_{n-1}}$, Applied Mathematics and Computation, in press.

[4] H. El-Metwally, E. A. Grove, and G. Ladas, A Global Convergence Result with Applications to Periodic Solutions, J. Math. Anal. Appl., 245 (2000), 161-170.

[5] H. El-Metwally, E. A. Grove, G. Ladas and H. D. Voulov, On the Global Attractivity and the Periodic Character of some Difference Equations, J. Differ. Equations Appl., 7 (2001), 1-14.

[6] V. L. Kocic and G. Ladas, Global Behavior of Nonlinear Difference Equations of Higher Order with Applications, Kluwer Academic Publishers, Dordrecht, 1993.

[7] M. R. S. Kulenovic and G. Ladas, Dynamics of Second Order Rational Difference Equations with Open Problems and Conjectures, Chapman & Hall / CRC Press, 2001.

JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS,VOL.5,NO.2,115-131,2007,COPYRIGHT 2007 EUDOXUS PRESS,LLC

A Game Theoretical Coalition Model of International Fishing with Time Delay

Andrew Engel
SAS Institute Inc.
10188 Telesis Court, Suite 200
San Diego, CA 92121, U. S. A.
andrew.engel@sas.com

Ferenc Szidarovszky
Systems and Industrial Engineering Department,
University of Arizona,
Tucson, Arizona 85721-0020, U. S. A.
szidar@sie.arizona.edu

Carl Chiarella
School of Finance and Economics,
University of Technology Sydney,
P.O. Box 123, Broadway, NSW 2007,
Australia
carl.chiarella@uts.edu.au

Abstract

The oligopoly model of international fishing of Szidarovszky and Okuguchi [7] where the harvesting countries form a coalition is revisited with the additional assumption that there is a time lag in obtaining and implementing information on the fish stock. The introduction of continuously distributed time lags results in a special Volterra-type integro-differential equation. Since it is equivalent to a system of nonlinear ordinary differential equations, linearization and

standard techniques are used to examine the local asymptotic behavior of the equilibrium. Stability conditions are derived and in the case of instability special cyclic behavior is analyzed.

1 Introduction

In his earlier paper, Okuguchi [5] has analyzed international fishing of two countries under imperfect competition. His model and methodology were extended by Szidarovszky and Okuguchi [6] to the n-country case, where a general dynamic systems model was developed to describe the trajectory of the fish stock as a function of time. In addition a complete equilibrium and stability analysis was performed. This model was further extended by Engel et al. [3] to include the inevitable time lags in obtaining and implementing information about the fish stock. Since the time delays are unknown, they were considered as random variables, and therefore this model was based on continuously distributed time lags.

The competitive model of Szidarovszky and Okuguchi [6] was modified by Szidarovszky and Okuguchi [7] to consider the case where at each time period, the harvesting countries form a grand coalition and their total profit is maximized. In this paper, this model will be extended to include time lags.

The structure of this paper is as follows: section 2 describes the n-country fish-harvesting oligopoly when a grand coalition is formed and how information lags are modeled. Section 3 examines the local asymptotic properties of the steady state. Section 4 provides numerical examples of the birth of limit cycles. Section 5 concludes the paper.

2 The Mathematical Model

2.1 Optimal Harvesting in a Static Setting

Consider a sea region where a single species of fish is harvested by n countries, and assume that each country sells its harvest in r markets. In our model we do not assume that $r = n$. The markets can be the markets of the fishery countries as well as those of others. As in Szidarovszky and Okuguchi [6], let x_{ki} denote the amount of fish harvested by country k $(1 \leq k \leq n)$ and sold in market i $(1 \leq i \leq r)$. The total amount of fish harvested by country k is given by $X_k = \sum_{i=1}^{r} x_{ki}$ and the total amount of fish sold in market i is

$Y_i = \sum_{k=1}^{n} x_{ki}$. The price of fish in market i is assumed to be linear:

$$p_i = a_i - b_i Y_i \tag{1}$$

with $a_i, b_i > 0$. The fishing cost of country k is given by:

$$C_k = c_k + \gamma_k \frac{X_k^2}{X}, \tag{2}$$

where $c_k, \gamma_k > 0$, and X is the total level of fish stock. With the above function forms we follow the earlier studies mentioned before. With these price and cost functions, the profit of country k is given by:

$$\Pi_k = \sum_{i=1}^{r} p_i x_{ki} - \left(c_k + \gamma_k \frac{X_k^2}{X} \right). \tag{3}$$

We assume now that the fish harvesting countries form a grand coalition. Therefore the total profit of the coalition is:

$$\Pi = \sum_{k=1}^{n} \Pi_k = \sum_{i=1}^{r} \left(a_i - b_i Y_i \right) Y_i - \sum_{k=1}^{n} \left(c_k + \gamma_k \frac{X_k^2}{X} \right) \tag{4}$$

which is concave in the variables x_{ki}. Assuming interior optimum, the first order conditions for the coalition's profit maximization is given by:

$$\frac{\partial \Pi}{\partial x_{ki}} = a_i - 2b_i Y_i - 2\gamma_k \frac{X_k}{X} = 0$$

for all i and k. Adding these equations over i and simplifying gives

$$A - S - B\gamma_k \frac{X_k}{X} = 0,$$

where $S = \sum_{i=1}^{r} Y_i$ is the total fish harvest, $A = \sum_{i=1}^{r} \frac{a_i}{2b_i}$ and $B = \sum_{i=1}^{r} \frac{1}{b_i}$. Therefore the total fish harvest of country k is

$$X_k = \frac{X}{\gamma_k B} \left[A - S \right]. \tag{5}$$

By adding these equations for all values of k and simplifying gives the total fish harvest:

$$S = \frac{ACX}{CX + B} \tag{6}$$

with $C = \sum_{k=1}^{n} \frac{1}{\gamma_k}$. Note in the purely competitive case (see [6]), the total fish harvest, S_c is given by

$$S_c = \frac{2Af(X)}{1+f(X)} \tag{7}$$

with

$$f(X) = \sum_{k=1}^{n} \frac{1}{1+2B\frac{\gamma_k}{X}}. \tag{8}$$

Proposition 1 *The total fish harvest under the purely competitive case is always greater than in the case of a grand coalition i.e.* $S_c > S$ *for all* $X > 0$.

Proof

Let $u_k = \frac{X}{2B\gamma_k}$, and note $\sum_{k=1}^{n} u_k = \frac{XC}{2B}$, so equation (8) can be rewritten as

$$f(X) = \sum_{k=1}^{n} \frac{u_k}{u_k+1}. \tag{9}$$

Similarly,

$$\begin{aligned} \frac{CX}{CX+2B} = \frac{1}{1+\frac{2B}{CX}} = \frac{1}{1+\frac{1}{\sum_{k=1}^{n} u_k}} &= \frac{\sum_{k=1}^{n} u_k}{\sum_{k=1}^{n} u_k + 1} \\ &= \sum_{k=1}^{n} \frac{u_k}{\sum_{l=1}^{n} u_l + 1}. \end{aligned} \tag{10}$$

Since each term of the right hand side of (9) is larger than the corresponding term of that of equation (10) we have

$$f(X) > \frac{CX}{CX+2B}.$$

This inequality can be simplified to

$$2Af(X)(CX+B) > ACX\left[1+f(X)\right]$$

or

$$\frac{2Af(X)}{1+f(X)} > \frac{ACX}{CX+B},$$

which proves the assertion. ∎

2.2 Dynamic Harvesting

Now, it is assumed that in the absence of fishing, the fish stock changes according to the logistic law:

$$\dot{X} = (\alpha - \beta X)\,X,$$

where $\alpha, \beta > 0$. This law has been known to fit well with experimental data from many biological populations and is widely used in the fishery literature, including Clark [2]. Other more complicated reproduction laws can be also used in our analysis with minor modifications. In the presence of international commercial fishing under a grand coalition structure, the fish stock changes according to

$$\dot{X} = X\left(\alpha - \beta X - \frac{AC}{CX + B}\right), \tag{11}$$

where we assume that the grand coalition continues to take the optimal harvest in all time periods $t \geq 0$.

Let

$$G\,(X) = \frac{AC}{CX + B},$$

then $G(X)$ is strictly decreasing and convex. Therefore the number of positive equilibria of the dynamic system (11) is 0, 1, or 2, since they are the intercepts of linear function $\alpha - \beta X$ and the curve of $G(X)$.

As in Engel et al. [3], it is assumed that when each country makes its decision on the amount of fish to harvest, it has delayed information on the size of the fish stock. The delay is uncertain, so it may be conveniently modeled by using continuously distributed time lags similarly to the earlier studies of Chiarella and Szidarovszky [1] and Engel et al. [3].

Instead of knowing the real value of the fish stock, each country has some delayed estimate where the delay is random. To model this situation, the expectation of country k on the fish stock at time period t, $X^{Ek}(t)$ is formed according to

$$X^{Ek}\,(t) = \int_0^t w\,(t - s, T_k, m_k)\,X\,(s)\,ds, \tag{12}$$

which is a weighted average of all past values of the fish stocks. Here the weighting function w is selected as it was in the earlier cited literature:

$$w\,(t - s, T, m) = \begin{cases} \frac{1}{T} e^{-\frac{t-s}{T}} & \text{if } m = 0 \\ \frac{1}{m!}\left(\frac{m}{T}\right)^{m+1} (t - s)^m\, e^{-\frac{m(t-s)}{T}} & \text{if } m \geq 1 \end{cases}. \tag{13}$$

Here m is an integer and T is a positive parameter. When $m = 0$, the most weight is assigned to the most current data, and the weights are exponentially declining as the data becomes older. When $m \geq 1$, zero weight is given to the most current data, the weight increases to a maximum at $t - s = T$ and the weight declines exponentially for all earlier data. As m increases, the weighting function becomes more peaked around $t - s = T$, and if $m \to \infty$ or $T \to 0$, the function tends to the Dirac delta function centered at T or zero, respectively.

When the countries form a grand coalition, each country harvests the amount of fish according to equation (5), but uses its expectation of the fish stock, X^{Ek} instead of the true value, X. Notice that equations (5) and (6) imply that

$$X_k = g_k \left(X^{Ek} \right) = \frac{AX^{Ek}}{\gamma_k (CX^{Ek} + B)},$$

so dynamic system (11) is now modified to

$$\dot{X}\left(t\right) = X\left(t\right)\left(\alpha - \beta X\left(t\right)\right) - \sum_{k=1}^{n} g_k \left(X^{Ek}\left(t\right) \right). \tag{14}$$

The dynamic equation (14), with X^{Ek} defined according to equation (12), is a Volterra-type integro-differential equation, which is equivalent to a system of nonlinear ordinary differential equations as it has been shown for similar models in Chiarella and Szidarovszky [1]. Therefore standard methods known from the theory of ordinary differential equations can be used to investigate the asymptotic properties of the solution.

First, we will linearize equation (14) around the equilibrium $\overline{X}$, and then use standard techniques for analyzing the asymptotic properties of the trajectory.

The linearized version of equation (14) can be written as

$$\dot{X}_\delta\left(t\right) = \left(\alpha - 2\beta\overline{X}\right) X_\delta\left(t\right) - \sum_{k=1}^{n} g_k'\left(\overline{X}\right) \int_0^t w\left(t - s, T_k, m_k\right) X_\delta\left(s\right) ds, \tag{15}$$

where X_δ is the deviation of X from its equilibrium level. We look for the solution to this equation in the form $X_\delta\left(t\right) = e^{\lambda t} v$. Substituting this solution into equation (15) and letting $t \to \infty$ gives the characteristic equation

$$\lambda - \left(\alpha - 2\beta\overline{X}\right) + \sum_{k=1}^{n} g_k'\left(\overline{X}\right) \left(1 + \frac{\lambda T_k}{q_k}\right)^{-(m_k+1)} = 0. \tag{16}$$

Notice that this is equivalent to a polynomial equation, so there are finitely many eigenvalues. The asymptotic properties of the trajectory $X(t)$ depends on the locations of the eigenvalues. In the general case it seems impossible to obtain analytical results, so numerical methods are used to find the solutions of equation (16). In order to obtain analytical results that may give some insights into the dynamics of the system, we have to consider special cases.

3 Effect of Time Delay

3.1 The Symmetric Country Case

Assume symmetric countries and markets, where $T_1 = \cdots = T_n = T$, $a_1 = \cdots = a_r = a$, $b_1 = \cdots = b_r = b$ and $\gamma_1 = \cdots = \gamma_n = \gamma$. This implies $A = \frac{ra}{2b}$, $B = \frac{r}{b}$, $C = \frac{n}{\gamma}$ and

$$g(X) = \frac{AX}{\gamma\left(CX + B\right)},$$

where $g_1(X) = \cdots = g_n(X) = g(X)$, implying that

$$g'(X) = \frac{AB}{\gamma(CX + B)^2} > 0$$

for all X.

At any equilibrium $\overline{X}$, it is clear that $\alpha - \beta\overline{X} = G\left(\overline{X}\right)$, so $\alpha - 2\beta\overline{X} = 2G\left(\overline{X}\right) - \alpha$ which can be positive, negative, or even zero depending on the shape of function G and the location of the equilibrium.

Equation (16) can now be rewritten as polynomial equation

$$\left[\lambda - \left(\alpha - 2\beta\overline{X}\right)\right]\left(1 + \frac{\lambda T}{q}\right)^{m+1} + ng'\left(\overline{X}\right) = 0. \tag{17}$$

3.2 Eigenvalue Analysis

3.2.1 The Case $m = 0$

In the case of $m = 0$, equation (17) is the quadratic:

$$\left[\lambda - \left(\alpha - 2\beta\overline{X}\right)\right](1 + \lambda T) + ng'\left(\overline{X}\right) = 0,$$

that is

$$T\lambda^2 + \lambda\left(1 - T\left(\alpha - 2\beta\overline{X}\right)\right) + \left(ng'\left(\overline{X}\right) - \alpha + 2\beta\overline{X}\right) = 0. \tag{18}$$

We will consider four cases depending on the sign and magnitude of $\alpha - 2\beta\overline{X}$ and its relationship to $g'\left(\overline{X}\right)$.

Case 1: Assume first that $\alpha - 2\beta\overline{X} \leq 0$.

In this case, all coefficients are positive implying that the real parts of the roots are negative, and therefore the equilibrium is asymptotically stable.

Case 2: Assume next that $\alpha - 2\beta\overline{X} > ng'\left(\overline{X}\right)$.

Then the constant term is negative, showing the existence of two real roots, one is negative and the other is positive. Thus the equilibrium is unstable.

Case 3: Assume that $\alpha - 2\beta\overline{X} = ng'\left(\overline{X}\right)$.

Now the constant term is zero implying the existence of two real eigenvalues, at least one of them is zero. If $ng'\left(\overline{X}\right) = \frac{1}{T}$, then both roots are zero. If $ng'\left(\overline{X}\right) < \frac{1}{T}$, then the nonzero eigenvalue is negative. In these cases no conclusion about the stability of the equilibrium can be drawn. If $ng'\left(\overline{X}\right) > \frac{1}{T}$, then the nonzero eigenvalue is positive implying the instability of the equilibrium.

Case 4: Assume finally that $0 < \alpha - 2\beta\overline{X} < ng'\left(\overline{X}\right)$.

Then the constant term is positive. Now we have to consider three further cases.

(i) If in addition $\alpha - 2\beta\overline{X} < \frac{1}{T}$, then the coefficients are positive, and the equilibrium is asymptotically stable.

(ii) If $\alpha - 2\beta\overline{X} > \frac{1}{T}$, then the linear coefficient becomes negative implying the existence of roots with positive value or positive real parts with unstable equilibrium.

(iii) Finally, consider the borderline case if

$$\alpha - 2\beta\overline{X} = \frac{1}{T}. \tag{19}$$

Then there is a pair of pure conjugate complex roots raising the possibility of the birth of limit cycles. In order to apply the Hopf bifurcation theorem (see for example Guckenheimer and Holmes [4]) we select T as the bifurcation parameter and show that the real parts of the derivatives of the eigenvalues with respect to the bifurcation parameter are nonzero at the critical value. By differentiating equation (18) implicitly we have

$$\lambda^2 + 2T\lambda\dot{\lambda} + \dot{\lambda}\left(1 - T\left(\alpha - 2\beta\overline{X}\right)\right) - \lambda\left(\alpha - 2\beta\overline{X}\right) = 0, \tag{20}$$

which implies

$$\dot{\lambda} = \frac{\lambda\left(\alpha - 2\beta\overline{X}\right) - \lambda^2}{2T\lambda + \left[1 - T\left(\alpha - 2\beta\overline{X}\right)\right]}.$$

Notice that at the critical value, $T^* = \frac{1}{\alpha - 2\beta\overline{X}}$, so we have

$$\dot{\lambda}\Big|_{T=T^*} = \frac{1 - T^*\lambda}{2\,{T^*}^2}, \tag{21}$$

so that

$$Re\ \dot{\lambda}\Big|_{T=T^*} = \frac{1}{2\,{T^*}^2} \neq 0$$

implying the existence of a limit cycle around the equilibrium.

Hence we have the following result.

Proposition 2 *Assume that in the symmetric case $m = 0$. The equilibrium of system (14) is asymptotically stable if $\alpha - 2\beta\overline{X} < min\{ng^{'}(\overline{X}), \frac{1}{T}\}$. The equilibrium is unstable if either $\alpha - 2\beta\overline{X} > ng^{'}(\overline{X})$, or $\alpha - 2\beta\overline{X} = ng^{'}(\overline{X}) > \frac{1}{T}$, or $\frac{1}{T} < \alpha - 2\beta\overline{X} < ng^{'}(\overline{X})$. In the case of $\alpha - 2\beta\overline{X} = \frac{1}{T} < ng^{'}(\overline{X})$ there is a limit cycle around the equilibrium.*

3.2.2 The Case $m = 1$

When $m = 1$, the characteristic equation (17) becomes

$$\left[\lambda - \left(\alpha - 2\beta\overline{X}\right)\right]\left(1 + 2\lambda T + \lambda^2 T^2\right) + ng'\left(\overline{X}\right) = 0,$$

which is a cubic polynomial equation:

$$\begin{aligned} \lambda^3 T^2 + \lambda^2\left(2T - T^2\left(\alpha - 2\beta\overline{X}\right)\right) + \lambda\left(1 - 2T\left(\alpha - 2\beta\overline{X}\right)\right) & \\ + \left(ng'\left(\overline{X}\right) - \alpha + 2\beta\overline{X}\right) &= 0. \end{aligned} \tag{22}$$

The Routh-Hurwitz stability criterion implies that all roots have negative real parts if and only if all coefficients are positive and

$$\begin{aligned} \left(2T - T^2\left(\alpha - 2\beta\overline{X}\right)\right)\left(1 - 2T\left(\alpha - 2\beta\overline{X}\right)\right) & \\ -T^2\left(ng'\left(\overline{X}\right) - \alpha + 2\beta\overline{X}\right) &> 0 \end{aligned}$$

which is equivalent to the quadratic inequality of the form

$$2\left(\alpha-2\beta\overline{X}\right)^2T^2-T\left(ng'\left(\overline{X}\right)+4\left(\alpha-2\beta\overline{X}\right)\right)+2>0. \tag{23}$$

At this point, the structures of equations (22) and (23) are very similar to those studied earlier in Engel et al. [3].

We will consider three cases according to the sign of $\alpha-2\beta\overline{X}$.

Case 1: $\alpha-2\beta\overline{X}<0$.

In this case, all coefficients of (22) are necessarily positive. We need to consider three possibilities. If $\alpha-2\beta\overline{X}=-\frac{1}{8}ng'\left(\overline{X}\right)$, then (23) holds for all

$$T\neq-\frac{1}{\alpha-2\beta\overline{X}},$$

and the equilibrium is asymptotically stable. If $\alpha-2\beta\overline{X}<-\frac{1}{8}ng'\left(\overline{X}\right)$, then the left hand side of (23) has no real root, so (23) holds for all $T>0$. Therefore, the equilibrium is asymptotically stable. If $\alpha-2\beta\overline{X}>-\frac{1}{8}ng'\left(\overline{X}\right)$, then the left hand side of (23) has two real roots:

$$\begin{aligned} T_{1,2} \;=\; & \left[4\left(\alpha-2\beta\overline{X}\right)^2\right]^{-1}\cdot\left[ng'\left(\overline{X}\right)+4\left(\alpha-2\beta\overline{X}\right)\right. \\ & \left.\pm\sqrt{ng'\left(\overline{X}\right)\left[ng'\left(\overline{X}\right)+8\left(\alpha-2\beta\overline{X}\right)\right]}\right]. \end{aligned} \tag{24}$$

Clearly, both roots are positive and if $T<T_1$ or $T>T_2$, the equilibrium is asymptotically stable, but if $T_1<T<T_2$, the equilibrium is unstable.

Case 2: $\alpha-2\beta\overline{X}=0$.

Then all coefficients of (22) are positive, and from the simplified form of (23) we conclude that the equilibrium is asymptotically stable for $T<\frac{2}{ng'\left(\overline{X}\right)}$, and unstable if $T>\frac{2}{ng'\left(\overline{X}\right)}$

Case 3: $\alpha-2\beta\overline{X}>0$.

The coefficients of (22) are positive if and only if

$$\alpha-2\beta\overline{X}<min\left\{\frac{1}{2T},ng'\left(\overline{X}\right)\right\}. \tag{25}$$

The quadratic polynomial (23) has two roots, given by (24), both are positive. In addition, (24) and (25) imply that $T_1<\frac{1}{2(\alpha-2\beta\overline{X})}<T_2$. Therefore $\alpha-2\beta\overline{X}<\frac{1}{2T}$ can not occur as $T>T_2$, and if $T<T_1$, then it holds

necessarily. Therefore in this case the equilibrium is asymptotically stable if $T < T_1$, and unstable if either $T > T_1$ or $\alpha - 2\beta\overline{X} > ng'(\overline{X})$.

Having analyzed the conditions under which the equilibrium may lose asymptotical stability, we now consider the possibility that the loss of asymptotical stability is accompanied by the birth of limit cycle motion in the fish stock. When $\overline{X}$ loses asymptotic stability, the real part of an eigenvalue passes to zero from a negative value. A pure complex number $\lambda = ir$ solves equation (22) if and only if

$$\begin{aligned} -ir^3T^2 - r^2\left(2T - T^2\left(\alpha - 2\beta\overline{X}\right)\right) + ir\left(1 - 2T\left(\alpha - 2\beta\overline{X}\right)\right) & \\ + \left(ng'\left(\overline{X}\right) - \alpha + 2\beta\overline{X}\right) &= 0. \end{aligned}$$

Equating the real and imaginary parts to zero, we have

$$r^2 = \frac{1 - 2T\left(\alpha - 2\beta\overline{X}\right)}{T^2} = \frac{ng'\left(\overline{X}\right) - \alpha + 2\beta\overline{X}}{2T - T^2\left(\alpha - 2\beta\overline{X}\right)}. \tag{26}$$

Since $T^2 > 0$, real nonzero r exists only if

$$1 - 2T\left(\alpha - 2\beta\overline{X}\right) > 0$$

and (23) is satisfied with equality:

$$2\left(\alpha - 2\beta\overline{X}\right)^2 T^2 - T\left(ng'\left(\overline{X}\right) + 4\left(\alpha - 2\beta\overline{X}\right)\right) + 2 = 0. \tag{27}$$

We consider two cases depending on whether $\alpha - 2\beta\overline{X}$ is zero or not.

Case 1: Assume first that $\alpha - 2\beta\overline{X} = 0$.

Then equation (27) is linear implying that

$$T^* = \frac{2}{ng'\left(\overline{X}\right)}$$

is the critical value. Differentiating equation (22) with respect to T implicitly we have

$$3\lambda^2\dot{\lambda}T^2 + 2\lambda^3T + 4\lambda\dot{\lambda}T + 2\lambda^2 + \dot{\lambda} = 0$$

implying that

$$\dot{\lambda} = -\frac{2\lambda^3T + 2\lambda^2}{3\lambda^2T^2 + 4\lambda T + 1} \tag{28}$$

Since $\lambda = ir$, the real part of this derivative is

$$Re\ \dot{\lambda}\Big|_{\lambda=ir} = \frac{A}{B} > 0$$

with

$$A = 2r^2\left(r^2T^2 + 1\right) > 0$$

and

$$B = \left(-3r^2T^2 + 1\right)^2 + (4rT)^2 > 0.$$

Therefore a limit cycle is born around the equilibrium.

Case 2: Assume next that $\alpha - 2\beta\overline{X} \neq 0$.

A positive solution for T exists only if the discriminant is nonnegative and the linear coefficient is negative:

$$\alpha - 2\beta\overline{X} \geq -\frac{1}{8}ng'\left(\overline{X}\right). \tag{29}$$

Let T_1 and T_2 ($T_2 \geq T_1$) denote the roots. If the bifurcation parameter is again selected as T, then T_1 and T_2 are the critical values. Differentiate equation (22) with respect to T implicitly to obtain

$$\begin{aligned} 3\lambda^2\dot{\lambda}T^2 + \lambda^3 2T + 2\lambda\dot{\lambda}\left(2T - T^2\left(\alpha - 2\beta\overline{X}\right)\right) & \\ +\lambda^2\left(2 - 2T\left(\alpha - 2\beta\overline{X}\right)\right) + \dot{\lambda}\left(1 - 2T\left(\alpha - 2\beta\overline{X}\right)\right) & \\ +\lambda\left(-2\left(\alpha - 2\beta\overline{X}\right)\right) &= 0 \end{aligned}$$

implying that

$$\dot{\lambda} = -\frac{2T\lambda^3 + \lambda^2\left(2 - 2T\left(\alpha - 2\beta\overline{X}\right)\right) - 2\lambda\left(\alpha - 2\beta\overline{X}\right)}{3\lambda^2T^2 + 2\lambda\left(2T - T^2\left(\alpha - 2\beta\overline{X}\right)\right) + \left(1 - 2T\left(\alpha - 2\beta\overline{X}\right)\right)}. \tag{30}$$

At $\lambda = ir$, the real part of this derivative is the following:

$$Re\ \dot{\lambda}\Big|_{\lambda=ir} = \frac{A}{B}, \tag{31}$$

with

$$\begin{aligned} B &= \left[-3r^2T^2 + 1 - 2T\left(\alpha - 2\beta\overline{X}\right)\right]^2 + \left[2r\left(2T - T^2\left(\alpha - 2\beta\overline{X}\right)\right)\right]^2 \\ &= \left[-2r^2T^2\right]^2 + \left[3rT + r^3T^3\right]^2 > 0 \end{aligned}$$

and similarly

$$A = r^2 \left(-r^4 T^4 + 2r^2 T^2 + 3 \right)$$

where we used relation (26). Notice that A is nonzero if $r^2 T^2 \neq 3$. In this case, there is a limit cycle as a consequence of the Hopf bifurcation theorem. Notice that (26) implies that $r^2 T^2 = 3$ if and only if

$$T \left(\alpha - 2\beta \overline{X} \right) = -1,$$

which is equivalent to equation

$$\alpha - 2\beta \overline{X} = -\frac{1}{8} n g' \left(\overline{X} \right), \tag{32}$$

which is (29) with equality. Therefore, if $\alpha - 2\beta \overline{X} > -\frac{1}{8} n g' \left(\overline{X} \right)$, then there is a limit cycle around the critical values, $T = T_1$ and $T = T_2$, otherwise the existence of a limit cycle is not guaranteed.

Thus we have the following result.

Proposition 3 *Assume that in the symmetric case $m = 1$. The equilibrium of system (14) is asymptotically stable if either*

$$\alpha - 2\beta \overline{X} < -\frac{1}{8} n g^{'} (\overline{X}),$$

or

$$\alpha - 2\beta \overline{X} = -\frac{1}{8} n g^{'} (\overline{X}) \neq -\frac{1}{T},$$

or

$$0 > \alpha - 2\beta \overline{X} > -\frac{1}{8} n g^{'} (\overline{X}) \quad \text{and} \quad T < T_1 \quad \text{or} \quad T > T_2$$

(where T_1 and T_2 are given in equation (24)),

or

$$\alpha - 2\beta \bar{X} = 0 \quad \text{and} \quad T < \frac{2}{n g^{'} (\bar{X})},$$

or

$$0 < \alpha - 2\beta \overline{X} < n g^{'} (\overline{X}) \quad \text{and} \quad T < T_1.$$

The equilibrium is unstable if either

$$0 > \alpha - 2\beta \overline{X} > -\frac{1}{8} n g^{'} (\overline{X}) \quad \text{and} \quad T_1 < T < T_2,$$

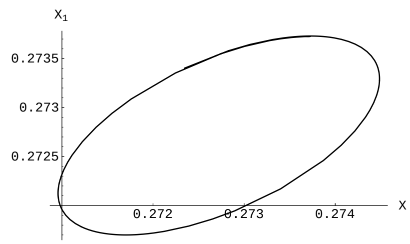


Figure 1: Birth of a Limit Cycle for $m = 0$

or

$$\alpha - 2\beta\overline{X} = 0 \quad \textit{and} \quad T > \frac{2}{ng^{'}(\overline{X})},$$

or

$$0 < \alpha - 2\beta\overline{X} < ng^{'}(\overline{X}) \quad \textit{and} \quad T > T_1,$$

or

$$\alpha - 2\beta\overline{X} > ng^{'}(\overline{X}).$$

There is limit cycle around the equilibrium if either

$$\alpha - 2\beta\overline{X} = 0 \quad \textit{and} \quad T = \frac{2}{ng^{'}(\overline{X})},$$

or

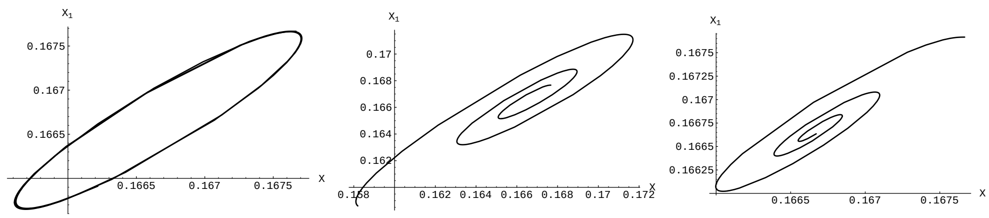
$$0 > \alpha - 2\beta\overline{X} > -\frac{1}{8}ng^{'}(\overline{X}) \quad \textit{and} \quad T \in \{T_1, T_2\},$$

or

$$0 < \alpha - 2\beta\overline{X} < ng^{'}(\bar{X}) \quad \textit{and} \quad T = T_1.$$

4 Numerical Examples

First, we illustrate the case of $m = 0$. We choose $n = r = 2$, $a = 1$, $b = 3$, $\gamma = 1$, and $\alpha = \beta = 1$. These parameters yield an equilibrium at $\overline{X} = 0.272727$. The critical value of T is 9.99991. Figure 1 shows the birth of a limit cycle with this value of T.



(a) Birth of a Limit Cycle

(b) An Expanding Limit Cycle

(c) A Shrinking Limit Cycle

Figure 2: Limit Cycles for $m = 1$

Next, we illustrate the case of $m = 1$. Here, we select $n = r = 2$, $a = 1$, $b = 3$, $\gamma = 1$, $\alpha = 0.94$ and $\beta = 1.64$. This choice of parameters yields an equilibrium at $\overline{X} = 0.166667$. There are two critical values: $T_1^* = 1.21909$ and $T_2^* = 5.30202$. Figure 2(a) shows the birth of a limit cycle for the critical value T_1^*. Figure 2(b) shows a expanding limit cycle for $T = 1.31909$ and figure 2(c) shows a shrinking limit cycle for $T = 1.11909$.

5 Conclusions

In this paper an international fishing model has been analyzed, when the fishing countries form a grand coalition and therefore at each time period their total profit is maximized. This situation was modeled by a modified version of the well known logistic law. We have shown that there are at most two steady states. Continuously distributed time lag was then assumed in obtaining and implementing information on the fish stock by the fishing countries. The corresponding dynamics of the fish stock resulted in a Volterra-type integro-differential equation. The asymptotical properties of the fish stock trajectory was examined by using linearization. In the case of symmetric countries and markets, we were able to derive a simple expression of the characteristic polynomial. A complete stability analysis was given based on the eigenvalues, and the possibility of the birth of limit cycles was

explored. Numerical examples illustrated such cases. For the nonsymmetric case and for larger values of m, numerical simulation is needed. Finally we mention that the same model and methodology can be used in investigating the long term effect of fishing in any given region or country, when the fishing firms play the same role as the countries in our case.

References

[1] C. Chiarella and F. Szidarovszky. The Birth of Limit Cycles in Nonlinear Oligopolies with Continuously Distributed Information Lags. In M. Dror, P. L'Ecuyer, and F. Szidarovszky, editors, *Modeling Uncertainty*, pages 249–268. Kluwer Academic Publishers., Dordrecht, 2001.

[2] C. W. Clark. *Mathematical Bioeconomics*. John Wiley and Sons, New York/London, 1976.

[3] A. Engel, C. Chiarella, and F. Szidarovszky. A Game Theoretical Model of International Fishing with Time Delay. In *Proceedings of the 2001 IEEE International Conference on Systems, Man and Cybernetics*, pages 2658–2663, Tucson, Arizona, October 7-10, 2001, 2001.

[4] J. Guckenheimer and P. Holmes. *Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields*. Springer - Verlag, New York, 1983.

[5] K. Okuguchi. Long-run Fish Stock and Imperfectly Competitive Commercial Fishing. *Keio Economic Studies*, 35:9–17, 1998.

[6] F. Szidarovszky and K. Okuguchi. An Oligopoly Model of Commercial Fishing. *Seoul Journal of Economics*, 11(3):321–330, 1998.

[7] F. Szidarovszky and K. Okuguchi. A Dynamic Model of International Fishing. *Seoul Journal of Economics*, 13(4):471–476, 2000.

JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS,VOL.5,NO.2,133-142,2007,COPYRIGHT 2007 EUDOXUS PRESS,LLC

On Starting Values for Parameters of Nonlinear Growth Functions with Application in Animal Science

Esdras Penêdo de Carvalho
Departamento de Matemática - Universidade Estadual de Maringá
epcarvalho@uem.br 87020-900 Maringá - PR, Brazil

Kleber Régis Santoro
Departamento de Zootecnia - Universidade Federal Rural de Pernambuco
ksantoro@ufrpe.br 55292-901 Garanhuns - PE, Brazil

Abstract

This paper is concerned with finding starting values for the parameters of von Bertalanffy, Brody, Gompertz and Logistic growth functions. A method is established by using the Walford plot and it is illustrated with examples from beef cattle. The method works with non-equally time-spaced data or non-monotone data, as in the case of animal growth with dry period.

Keywords: Nonlinear regression, growth function, scientific computation, animal science.

1 Introduction

The analysis of growth patterns in animals and plants is a standard procedure in research fields such as animal science, forestry and fishery management, among others. In many cases, exponential functions are used for modeling the nonlinear relation between weight (or length) and age, as for instance, in genetic improvement of dairy and beef cattle (Brown et al [3] and Perotto et al [13]). In cattle research, four growth functions play an important role,

namely, von Bertalanffy [1], Brody [2], Gompertz [18] and Logistic [12]. After re-calibrating the response variables (weights), all the above cited functions can be written in the form

$$w = A(1 - Be^{-Kt}),$$

where t denotes animal age, $w = w(t)$ is a suitable re-scaling of the animal weight at age t, and A, B, K are parameters of the model. As a matter of fact, fitting these curves to experimental data, iterative methods of nonlinear regression such as Gauss-Newton or Levenberg-Marquardt [10] are used. It turns out that selecting starting values for the parameters A, B, K is an important step in this process. Giving bad starting values may slow down the convergence speed and sometimes may cause the iterative process to fail or to converge to unwanted results.

There are some works dedicated to the computation of starting values for parameters of nonlinear curves. Typically there are very technical and often avoided by non-mathematicians (e.g. Causton [4] or Nelder [12]). Some other works suggest finding starting values in an empirical way, based on early knowledge or on similar experiments found in the literature (Santoro et al [15]).

The goal of this paper is to provide a mathematically consistent method of finding starting values for the parameters of the above four growth functions. The method can handle non-equally time-spaced or non-increasing data, and it is sufficiently simple to be used by a broad class of scientists, specially in animal science. Examples are provided with data obtained from Brazilian zebu cattle database.

2 Description of the Method

The method is based on the Walford plot (Walford [17]), widely used in fishery biology to estimate asymptotic length of several aquatic animals. The Brody function, also known as von Bertalanffy (1938) in fishery biology, will be considered first, since the other three functions can be reduced to the former by re-scaling w_i data.

The following conventions will be used throughout this section.

Table 1: Definition of models and variables

BRD	Brody function: $w_i = A\left(1 - Be^{-Kt_i}\right)$
VBF	von Bertalanffy function: $w_i = A\left(1 - Be^{-Kt_i}\right)^3$
LOG	Logistic function: $w_i = A\left(1 + Be^{-Kt_i}\right)^{-1}$
GMP	Gompertz function: $w_i = Ae^{-Be^{-Kt_i}}$
t_i	A serie of ages (days, months or years)
w_i	Weight at age t_i (maybe length)
N	Number of data $(1 \le i \le N)$
Δt	Average of $\Delta t_i = t_{i+1} - t_i$
K	growth rate or maturity coefficient
A	asymptotic weight or length
B	integration constant related to w_0

The Brody function – From the Brody function

$$w_i = A\left(1 - Be^{-Kt_i}\right),$$

a simple relation between w_i and w_{i+1} can be obtained. Indeed, one has

$$w_{i+1} = A(1 - e^{-K\Delta t_i}) + e^{-K\Delta t_i} w_i.$$

If t_i are equally spaced, then $\Delta t_i = \Delta t$ is a constant for every i. Otherwise, define Δt as the average of all Δt_i, that is,

$$\Delta t = \frac{1}{N-1} \sum_{i=1}^{N-1} (t_{i+1} - t_i).$$

It follows that the approximation

$$w_{i+1} \approx A(1 - e^{-K\Delta t}) + e^{-K\Delta t} w_i,$$

can be assumed, and therefore the parameters K and A are approximated by fitting a line $y = a + bx$ to the plot w_i versus w_{i+1}. Of course, x represents w_i and y represents w_{i+1}. After obtaining a and b by least squares, one takes as starting values

$$K = -\frac{\ln b}{\Delta t} \quad \text{and} \quad A = \frac{a}{1 - e^{-K\Delta t}}.$$

Since B can be deduced from Brody function as

$$B = e^{Kt_i}\left(1 - \frac{w_i}{A}\right), \quad i = 1, 2, \cdots, N,$$

its value can be approximated as the average of B, for $1 \le i \le N$. A summary is presented in the Table 2.

The von Bertalanffy and Gompertz functions – These two functions can be studied as above by a suitable re-scaling of w_i data. Indeed, re-scaling w_i to $y_i = \sqrt[3]{w_i}$ transforms VBF to

$$y_i = \sqrt[3]{A}\left(1 - Be^{-Kt_i}\right),$$

which is a Brody type function with y_i in the place of w_i and $\sqrt[3]{A}$ in the place of A. Then, fitting a line $y_{i+1} = a + by_i$ to y_i against y_{i+1} by least squares, similar results hold:

$$K = -\frac{\ln b}{\Delta t}, \quad A = \frac{a^3}{(1 - e^{-K\Delta t})^3}, \quad B = \frac{1}{N}\sum_{i=1}^{N} e^{Kt_i}\left(1 - \sqrt[3]{\frac{w_i}{A}}\right).$$

Likewise, re-scaling w_i to $\ln w_i$ transforms GMP to

$$y_i = (\ln A)\left(1 - \frac{B}{\ln A}\, e^{-Kt_i}\right).$$

This is also a Brody type function with y_i in the place of w_i, $\ln A$ in the place of A and $B/\ln A$ in the place of B. Then fitting a line $y_i = a + by_{i+1}$ to the plot $\ln w_i$ versus $\ln w_{i+1}$, similar results also hold:

$$K = -\frac{\ln b}{\Delta t}, \quad A = e^{a/(1-e^{-K\Delta t})}, \quad B = \frac{1}{N}\sum_{i=1}^{N} e^{Kt_i}\ln\left(\frac{A}{w_i}\right).$$

See summary in Table 2.

The Logistic function – The Logistic function can be reduced to Brody function by transforming w_i to $y_i = 1/w_i$. This gives

$$y_i = A^{-1}\left(1 - (-B)e^{-Kt_i}\right).$$

Then with a similar argument, the proposed starting values are:

$$K = -\frac{\ln b}{\Delta t}, \quad A = \frac{(1 - e^{-K\Delta t})}{a}, \quad B = \frac{1}{N}\sum_{i=1}^{N} e^{Kt_i}\left(\frac{A}{w_i} - 1\right).$$

However, due to numerical instability, this argument will only work for data with a clear sigmoidal shape. On the other hand, the Logistic function can be easily linearized when the asymptotic value A is known. Indeed, the logistic function can be written as

$$\frac{A}{w_i} = 1 + Be^{-Kt_i},$$

and therefore

$$\ln\left(\frac{A}{w_i} - 1\right) = \ln B - Kt_i.$$

Then if the parameter A is given, a line $y = a + b\,t$ can be fitted to t_i versus $y_i = \ln(A/w_i - 1)$. Usually one takes A as 110% of the maximum value of w_i. Hence parameters A, B and K can be approximated by

$$A = 1.1 \times \max\{w_i\} \qquad B = e^a \qquad \text{and} \qquad K = -b.$$

This procedure is widely considered in the literature (e.g. Mathews [11]) and therefore it is recommended in the present paper. A summary is in Table 2.

Table 2: Starting values for the parameters A, B, K of the named functions. Given data t_i, w_i, first compute coefficients a, b with linear least square and then K, A, B.

Model	Fit line	K	A	B
BRD	$w_{i+1} = a + b\,w_i$	$-\dfrac{\ln b}{\Delta t}$	$\dfrac{a}{1 - e^{-K\Delta t}}$	$\dfrac{1}{N}\sum_{i=1}^{N} e^{Kt_i}\left(1 - \dfrac{w_i}{A}\right)$
VBF	$\sqrt[3]{w_{i+1}} = a + b\sqrt[3]{w_i}$	$-\dfrac{\ln b}{\Delta t}$	$\dfrac{a^3}{(1 - e^{-K\Delta t})^3}$	$\dfrac{1}{N}\sum_{i=1}^{N} e^{Kt_i}\left(1 - \sqrt[3]{\dfrac{w_i}{A}}\right)$
GMP	$\ln w_{i+1} = a + b\,\ln w_i$	$-\dfrac{\ln b}{\Delta t}$	$e^{a/(1-e^{-K\Delta t})}$	$\dfrac{1}{N}\sum_{i=1}^{N} e^{Kt_i}\ln\left(\dfrac{A}{w_i}\right)$
LOG	$\ln\left(\dfrac{A}{w_i} - 1\right) = a + b\,t_i$	$-b$	$1.1 \times \max\{w_i\}$	e^a

3 Application to Animal Science

To illustrate the method described in Table 2, data from Brazilian Association of Zebu Breeders (ABCZ) are used. They refer to the average weights of 166 female zebu animals from Guzera and Nelore breeds, recorded between 1979 and 1995 in the state of Pernambuco, north-eastern of Brazil. The records were taken at birth and at, approximately, each 90 days up to two years.

Table 3: Average age and weight of 166 female zebus from ABCZ database.

Age (days)	t_i	0	85	174	264	357	448	538	630	723
Weight(kg)	w_i	30	86	143	190	224	253	284	313	339

For sake of understanding, an outline of the computations used for VBF function following Table 2 is presented now. First transform w_i to $y_i = \sqrt[3]{w_i}$, that is,

$$y = [3.1072,\ 4.4140,\ 5.2293,\ 5.7489,\ 6.0732,\ 6.3247,\ 6.5731,\ 6.7897].$$

Then fit a line $y_{i+1} = a + by_i$ with least square method to get

$$a = 2.2061 \quad \text{and} \quad b = 0.6886.$$

Since the average time-interval is

$$\Delta t = \frac{1}{8}\{85 + 89 + 90 + 93 + 91 + 90 + 92 + 93\} = 90,$$

it follows (using Table 2) that

$$K = -\frac{\ln 0.6886}{90} = 0.0041455, \qquad A = \frac{(2.2061)^3}{(1 - e^{-0.0041455 \times 90})^3} = 3.5551,$$

$$B = \frac{1}{9}\sum_{1}^{9}\left[e^{-0.0041455 \times t_i} \times \left(1 - \sqrt[3]{\frac{w_i}{3.5551}}\right)\right] = 0.56280.$$

The precision of the parameters computed using Table 2 are estimated by comparing them with the converged values using Levenberg-Marquardt method. Standard deviation of the residuals

$$\text{SDR} = \sqrt{\left(\frac{1}{N-1}\sum_{i=1}^{N}(r_i - \overline{r})^2,\right)}$$

of fitting starting curves to data are shown in the Table 4. Here, starting curves mean the growth curve using the corresponding starting parameters. Compare these values with the converged values in Table 5. Graphics comparing starting curves and converged curves are presented in the Figure 1.

Table 4: Starting parameters computed using Table 2 and data from Table 3.

Model	A	B	K	SDR
BRD	503.6	0.9402	0.0015	2.963
VBF	355.5	0.5628	0.0041	6.136
GMP	321.0	1.8026	0.0061	9.919
LOG	378.8	6.2002	0.0058	8.816

Table 5: Converged parameters through Levenberg-Marquardt method using data from Table 3 and starting values given in Table 4.

Model	A	B	K	SDR
BRD	471.1	0.9338	0.0016	1.999
VBF	381.0	0.5279	0.0034	3.643
GMP	364.0	2.0975	0.0042	4.764
LOG	339.7	5.1875	0.0066	8.063

4 Conclusions

Analyzing comparatively tables 4 and 5 and figures 1(a) to 1(d), it is clear that starting values computed using Table 2 are very close to the converged values when data have the shape of the corresponding curves. Now, limitations do exist in some cases. Methods based on Walford plot rely, in fact, on uniformly time-spaced data, because the need of Δt. But since the main point is computing starting (approximate) values for parameters, an approximate value of Δt will works properly. However, if Δt_i have a large standard deviation, the method developed in Table 2 may give bad results. A different method of fitting curves with non-uniformly spaced data is presented by Gulland and Holt [8]. It works well in fishery biology but seems to be unsuitable

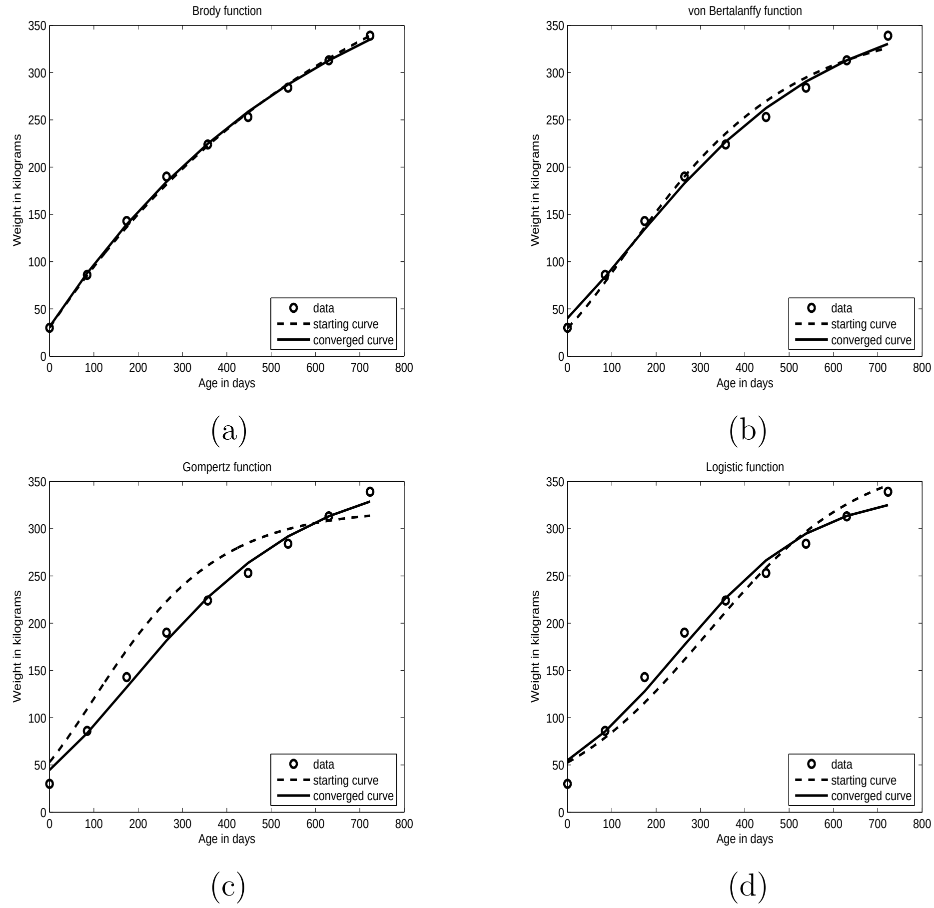


Figure 1: Comparing starting curves with the fitted curves for: (a) Brody, (b) von Bertalanffy, (c) Gompertz and (d) Logistic functions.

for analyzing grass-fed cattle data. This happens because the animals may lose weight during the dry period.

Finally, it is worth noting that the use of empirical arguments for estimating starting values of the parameters A, B, K is still very common in the literature. Indeed, they have shown that this strategy does work in most cases. Then the present review offers an alternative mathematical framework for this task, which may be very useful when considering situations where previous empirical evidences are not known. Serious experimental researches always take advantages of mathematical deductions complemented with empirical knowledge.

Acknowledgement. The authors are grateful to the Brazilian agency CNPq through project Edital Universal for financial support. They also thank Professor To Fu Ma for his comments and suggestions which improved an early version of this paper.

References

[1] L. von Bertalanffy, A quantitative theory of organic growth, Human Biol. 10, 1938, 181-213.

[2] S. Brody, Bioenergetics and Growth, Rheinhold Pub. Corp., New York, 1945, 645pp.

[3] J. E. Brown, H. A. Fitzhugh Jr, T. C. Cartwright, A comparison of nonlinear models for describing weight-age relationships in cattle, Journal of Animal Science 42, 1976, 810-818.

[4] D. R. Causton, A computer program for fitting the richards function, Biometrics 25, 1969, 401-409.

[5] C. W. Clark, Mathematical Bioeconomics, Second Edition, John Wiley and Sons, New York, 1990.

[6] S. G. Dastidar, Gompertz: A Scilab program for estimating Gompertz curve using Gauss-Newton method of least squares, J. Statistical Software 15, (2006), 1-12.

[7] H. A. Fitzhugh Jr, Analysis of growth curves ans strategies for altering their shape, Journal of Animal Science 42, 1976, 1036-1051.

[8] J. A. Gulland and J. S. Holt, Estimation of growth parameter for data at unequal time intervals, Journal du Conseil International pour l'Éxploration de la Mer 25, 1959, 47-49.

[9] K. Levenberg, A method for the solution of certain problems in least-squares, Quarterly of Applied Mathematics 2, 1944, 164-168.

[10] D. Marquardt, An algorithm for least-squares estimation of nonlinear parameters, SIAM Journal of Applied Mathematics 11, 1963, 431-441.

[11] J. H. Mathews, Bounded population growth: a curve fitting lesson, Mathematics and Computer Education 26, 1992, 169-176.

[12] J. A. Nelder, The fitting of a generalization of the logistic curve, Biometrics 17, 1961, 89-94.

[13] D. Perotto, R. I. Cue and A. J. Lee, Comparison of nonlinear functions for describing the growth curve of three genotypes of dairy cattle, Canadian Journal of Animal Science 72, 1992, 773-782.

[14] F. J. Richards, A flexible growth function for empirical use, Journal of Experimental Botany 10, 1959, 290-300.

[15] K. R. Santoro, S. B. P. Barbosa, L. H. Albuquerque, E. S. Santos, Estimativas de parâmetros de curvas de crescimento de bovinos Zebu, criados no estado de Pernambuco, Brazilian Journal of Animal Science 34, 2005, 2262-2279.

[16] SAS Institute Inc. SAS/STAT User's guide, 2003.

[17] L. A. Walford, New graphic method of describing the growth of animals, Bulletin of the U. S. Bureau of Fisheries 46, 1946, 633-641.

[18] C. P. Winsor, The Gompertz curve as a growth curve, Proceedings of the National Academy of Science 18, 1932, 1-8.

JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS,VOL.5,NO.2,143-152,2007,COPYRIGHT 2007 EUDOXUS PRESS,LLC

Numerical Solution of a Nonlinear Coupled Advection-Diffusion Equation

Robert R. Ferdinand
Mathematics Department
East Central University
Ada, OK 74820, U.S.A.
Email: robert.ferdinand@ecok.edu

March 13, 2006

Abstract

In this paper the author presents a coupled system of two nonlinear, advection-diffusion equations. Similar systems of equations have been used to describe the dynamics of contaminant in groundwater as it flows through cracks in a porous rock matrix and gets consequently absorbed into it. An explicit finite difference scheme is used to numerically approximate the solution of this mathematical model. Convergence of the numerical approximants, thereby obtained, to a unique continuous solution is proved and then the order of convergence of this numerical scheme is illustrated via a computational experiment. Graphical results are displayed to illustrate the order of convergence and finally, future research work to be done is discussed.

Mathematics Subject Classification Numbers: 35K55, 35K57, 65N06, 65N12, 65Y20

Key Words: advection, coupled, diffusion, finite difference, numerical, order of convergence

1 Introduction

In [3, 6] the authors present a model that describes the dynamics of contaminant present in groundwater as it flows through a system of parallel cracks and gets consequently absorbed into the surrounding porous rocky medium. The reader may refer to [3, 6] for a further detailed and illustrated description of this physical phenomenon.

In [3], the authors use the Bromwich Inversion Integral and residue calculus to obtain an exact solution in the form of a single, infinite integral. This solution is much simpler than the one obtained in [6], which involves the use of double and triple integrals.

The mathematical model studied in this manuscript is fairly similar to the one described in [3, 6] but also vastly different since it contains nonlinear advection and diffusion terms.

This model, represented by a nonlinear coupled system of advection-diffusion PDEs, hence presents itself as follows:

$$\begin{cases} C_t + [a(C)]_z = [b(C, C_z)]_z - \lambda C + \beta M_x(t, 0, z) \\ M_t = [e(M, M_x)]_x - \lambda M \end{cases} \tag{1 − 2}$$

where $C(t, z)$ represents the contaminant concentration in the liquid flowing through a fracture in the porous rock matrix, while $M(t, x, z)$ is the contaminant concentration in the liquid as it diffuses into the rock matrix while flowing through the fracture. The variable z represents distance along the fracture, x the perpendicular distance into the rock matrix and t the time. Initial and boundary conditions for $C(t, z)$ are given by

$$\begin{cases} C(0, z) = \alpha(z) \\ C(t, 0) = \gamma(t) \\ C(t, z_{\max}) = 0, \end{cases} \tag{3}$$

while initial and boundary conditions for $M(t, x, z)$ are given by

$$\begin{cases} M(0, x, z) = \eta(x, z) \\ M(t, 0, z) = \rho\, C(t, z) \\ M(t, x_{\max}, z) = 0. \end{cases} \tag{4}$$

Further, the following regularity conditions, necessary for convergence, are also imposed on the model parameters:

(J1) Nonlinear advection term a in Eq.(1) is uniformly Lipschitz in C.

(J2) Nonlinear diffusion term b in Eq.(1) is uniformly Lipschitz in C and C_z.

(J3) Nonlinear diffusion term e in Eq.(2) is uniformly Lipschitz in M and M_x.

(J4) Functions α, γ and η are bounded and measurable on $[0, \infty)$.

Although more numerically stable implicit finite difference schemes are used by the authors in [1, 2] to numerically approximate solutions of the mathematical models therein, the author here uses an explicit finite difference scheme, similar to the one used in [4], to obtain a numerical solution of Eqs.(1-2).

This paper is therefore organized in the following fashion. In the next section, the explicit finite difference solution approximating scheme is described while convergence of the solution approximants to a unique solution is analytically proved in section 3. A computational experiment that illustrates the order of convergence of this numerical method is presented in section 4, while future research work to be done with this model is finally discussed in section 5.

2 Finite Difference Solution Approximating Scheme

The following uniform mesh sizes are used for the discretization of the t, x and z axes, respectively:

$$\Delta t = \frac{T_{\max}}{Q}, \qquad \Delta x = \frac{x_{\max}}{N} \qquad \text{and} \qquad \Delta z = \frac{z_{\max}}{R},$$

which leads to taking the mesh points on the t, x and z axes, respectively, as follows:

$$t_k = k\ \Delta t, \qquad x_j = j\ \Delta x \qquad \text{and} \qquad z_l = l\ \Delta z,$$

for $k = 0, \cdots, Q$, $j = 0, \cdots, N$ and $l = 0, \cdots, R$.

The notation C_l^k and $M_{j,l}^k$ is used to represent $C\left(t_k, z_l\right)$ and $M\left(t_k, x_j, z_l\right)$, respectively, while similar notation is used to represent the other model functions and parameters. Hence the following explicit finite difference scheme, Eqs.(5-8) below, is used to numerically approximate the solution of the coupled system, Eqs.(1-4):

$$\begin{aligned} &\frac{C_l^{k+1} - C_l^k}{\Delta t} + \left[\frac{a\left(C_{l+1}^k\right) - a\left(C_l^k\right)}{\Delta z}\right] \\ &= \left[\frac{b\left(C_l^k, \dfrac{C_{l+1}^k - C_l^k}{\Delta z}\right) - b\left(C_{l-1}^k, \dfrac{C_l^k - C_{l-1}^k}{\Delta z}\right)}{\Delta z}\right] \\ &-\lambda C_l^k + \beta\left(\frac{M_{1,l}^k - \rho\ C_l^k}{\Delta x}\right), \end{aligned} \tag{5}$$

for $l = 1, \cdots, R-1$ with Eq.(3) giving

$$C_l^0 = \alpha_l, \qquad C_0^k = \gamma^k \qquad \text{and} \qquad C_R^k = 0. \tag{6}$$

Further,

$$\frac{M_{j,l}^{k+1} - M_{j,l}^k}{\Delta t} = \left[\frac{e\left(M_{j,l}^k, \dfrac{M_{j+1,l}^k - M_{j,l}^k}{\Delta x}\right) - e\left(M_{j-1,l}^k, \dfrac{M_{j,l}^k - M_{j-1,l}^k}{\Delta x}\right)}{\Delta x}\right] - \lambda\ M_{j,l}^k, \tag{7}$$

for $j = 1, \cdots, N-1$ and $l = 0, \cdots, R$ with Eq.(4) yielding

$$M_{j,l}^0 = \eta_{j,l}, \qquad M_{0,l}^k = \rho\ C_l^k \qquad \text{and} \qquad M_{N,l}^k = 0. \tag{8}$$

Now, let

$$\mu_1 = \frac{\Delta t}{\Delta z} \qquad \text{and} \qquad \mu_2 = \frac{\Delta t}{\Delta x},$$

and calculate the numerical solution $\left(C_l^k, M_{j,l}^k\right)$ recursively using the known initial and boundary condition data from Eq.(6) and Eq.(8) and marching forward in time as follows:

$$\begin{aligned} C_l^{k+1} \;=\;& (1 - \lambda\, \Delta t)\, C_l^k + \mu_1 \left[a\left(C_l^k\right) - a\left(C_{l+1}^k\right)\right. \\ & +b\left(C_l^k, \frac{C_{l+1}^k - C_l^k}{\Delta z}\right) - b\left(C_{l-1}^k, \frac{C_l^k - C_{l-1}^k}{\Delta z}\right)\Bigg] \\ & +\beta\, \mu_2\, \left(M_{1,l}^k - \rho\, C_l^k\right) \end{aligned} \tag{9}$$

and

$$\begin{aligned} M_{j,l}^{k+1} \;=\;& (1 - \lambda\, \Delta t)\, M_{j,l}^k + \mu_2 \left[e\left(M_{j,l}^k, \frac{M_{j+1,l}^k - M_{j,l}^k}{\Delta x}\right)\right. \\ & \left. - \, e\left(M_{j-1,l}^k, \frac{M_{j,l}^k - M_{j-1,l}^k}{\Delta x}\right)\right]. \end{aligned} \tag{10}$$

Thus Eqs.(6,8-10) serve as the finished form of the explicit finite difference scheme which provides the numerical solution of this model and thereby concludes this section. In the next section, it will be shown that these solution approximants $\left(C_l^k, M_{j,l}^k\right)$ converge to the unique solution (C, M) of Eqs.(1-4).

3 Convergence and Uniqueness

This section is begun by stating and proving a theorem.

Theorem 1. Denote the numerical approximation of the solution of Eqs.(1-4) by $\left(C_l^k, M_{j,l}^k\right)$ and the exact solution by $\left(\hat{C}_l^k, \hat{M}_{j,l}^k\right)$ for $j = 0, \cdots, N$, $l = 0, \cdots, R$ and $k = 0, \cdots, Q$. Then

$$\left(C_l^k, M_{j,l}^k\right) \to \left(\hat{C}_l^k, \hat{M}_{j,l}^k\right)$$

as $\Delta t,\ \Delta x,\ \Delta z \to 0$.

Proof. Write Eqs.(9-10) in operator form as

$$C_l^{k+1} = A_1\left(C_{l-1}^k, C_l^k, C_{l+1}^k, M_{1,l}^k\right) + \omega_1\left[(\Delta t)^2 + \Delta t\,(\Delta x + \Delta z)\right] \tag{11}$$

and

$$M_{j,l}^{k+1} = A_2\left(M_{j-1,l}^k, M_{j,l}^k, M_{j+1,l}^k\right) + \omega_2\left[(\Delta t)^2 + \Delta x\ \Delta t\right], \tag{12}$$

respectively, since the finite difference scheme Eqs.(5-8) is $O(\Delta t, \Delta x, \Delta z)$ (see [5]). Further,

$$\hat{C}_l^{k+1} = A_1\left(\hat{C}_{l-1}^k, \hat{C}_l^k, \hat{C}_{l+1}^k, \hat{M}_{1,l}^k\right) \tag{13}$$

and

$$\hat{M}_{j,l}^{k+1} = A_2\left(\hat{M}_{j-1,l}^k, \hat{M}_{j,l}^k, \hat{M}_{j+1,l}^k\right). \tag{14}$$

Now let $\overline{C}_l^k = C_l^k - \hat{C}_l^k$ and $\overline{M}_{j,l}^k = M_{j,l}^k - \hat{M}_{j,l}^k$. Assuming C and $\hat{C}$ have identical initial and boundary conditions, one gets

$$\overline{C}_l^0 = \overline{C}_0^k = \overline{C}_R^k = 0, \tag{15}$$

while using techniques such as those used and described in [4, 5] and properties (J1)-(J4), subtract Eq.(13) from Eq.(11) and after elementary algebraic manipulations arrive at

$$\left|\overline{C}_l^{k+1}\right| \le \omega_3\left(\left|\overline{C}_{l-1}^k\right| + \left|\overline{C}_l^k\right| + \left|\overline{C}_{l+1}^k\right| + \left|\overline{M}_{1,l}^k\right|\right) + \omega_1\left[(\Delta t)^2 + \Delta t\,(\Delta x + \Delta z)\right], \tag{16}$$

for $l = 1, \cdots, R-1$.
Similarly, the same initial and boundary conditions for M and $\hat{M}$ lead to

$$\overline{M}_{0,l}^k = \rho\,\overline{C}_l^k, \qquad \overline{M}_{j,l}^0 = \overline{M}_{N,l}^k = 0, \tag{17}$$

while subtracting Eq.(14) from Eq.(12) and (J1)-(J4) yield

$$\left|\overline{M}_{j,l}^{k+1}\right| \le \omega_4\left(\left|\overline{M}_{j-1,l}^k\right| + \left|\overline{M}_{j,l}^k\right| + \left|\overline{M}_{j+1,l}^k\right|\right) + \omega_2\left[(\Delta t)^2 + \Delta x\ \Delta t\right] \tag{18}$$

for $j = 1, \cdots, N-1$ and $l = 0, \cdots, R$.
Now, define

$$\Gamma_{\overline{C}}^k = \max_{l=0,\cdots,R}\left|\overline{C}_l^k\right| \qquad \text{and} \qquad \Gamma_{\overline{M}}^k = \max_{j=0,\cdots,N\ ;\ l=0,\cdots,R}\left|\overline{M}_{j,l}^k\right|. \tag{19}$$

Then from Eqs.(15-16)

$$\Gamma_{\overline{C}}^{k+1} \le \omega_5\left(\Gamma_{\overline{C}}^k + \Gamma_{\overline{M}}^k\right) + \omega_1\left[(\Delta t)^2 + \Delta t\,(\Delta x + \Delta z)\right] \tag{20}$$

and from Eqs.(17-18)

$$\Gamma_{\overline{M}}^{k+1} \le \omega_6\ \left(\Gamma_{\overline{C}}^k + \Gamma_{\overline{M}}^k\right) + \omega_2\left[(\Delta t)^2 + \Delta x\ \Delta t\right]. \tag{21}$$

Thus, Eq.(20) leads to

$$\Gamma_{\frac{k}{C}} \leq \omega_7 \left(\Gamma_{\frac{0}{C}} + \Gamma_{\frac{0}{M}}\right) + \omega_8 \left[(\Delta t)^2 + \Delta t \left(\Delta x + \Delta z\right)\right], \tag{22}$$

while Eq.(21) gives

$$\Gamma_{\frac{k}{M}} \leq \omega_9 \left(\Gamma_{\frac{0}{C}} + \Gamma_{\frac{0}{M}}\right) + \omega_{10} \left[(\Delta t)^2 + \Delta x \ \Delta t\right]. \tag{23}$$

Passing the limit $\Delta t,\ \Delta x,\ \Delta z \to 0$ in Eqs.(22-23) one gets

$$\Gamma_{\frac{k}{C}} = \Gamma_{\frac{k}{M}} = 0$$

for $k = 0, \cdots, Q$. This leads to $C_l^k \to \hat{C}_l^k$ and $M_{j,l}^k \to \hat{M}_{j,l}^k$, which concludes the proof of this theorem. The next theorem proves uniqueness of this numerical solution obtained.

Theorem 2. The numerical solution arrived at in Theorem 1 above is unique.

Proof. Let $\left(\check{C}, \check{M}\right)$ and $\left(\acute{C}, \acute{M}\right)$ be two solutions of Eqs.(1-4) with both pairs having identical initial and boundary conditions. Also, let $\tilde{C} = \check{C} - \acute{C}$ and $\tilde{M} = \check{M} - \acute{M}$ and straight away get

$$\tilde{C}_l^0 = \tilde{C}_0^k = \tilde{C}_R^k = 0. \tag{24}$$

Further, just as in the proof of Theorem 1, Eq.(13) gives

$$\tilde{C}_l^{k+1} = A_1 \left(\check{C}_{l-1}^k, \check{C}_l^k, \check{C}_{l+1}^k, \check{M}_{1,l}^k\right) - A_1 \left(\acute{C}_{l-1}^k, \acute{C}_l^k, \acute{C}_{l+1}^k, \acute{M}_{1,l}^k\right),$$

which, using (J1)-(J4), leads to

$$\left|\tilde{C}_l^{k+1}\right| \leq \omega_{11} \left(\left|\tilde{C}_{l-1}^k\right| + \left|\tilde{C}_l^k\right| + \left|\tilde{C}_{l+1}^k\right| + \left|\tilde{M}_{1,l}^k\right|\right) \tag{25}$$

for $l = 1, \cdots, R-1$.
Next,

$$\tilde{M}_{0,l}^k = \rho \, \tilde{C}_l^k, \qquad \tilde{M}_{j,l}^0 = \tilde{M}_{N,l}^k = 0, \tag{26}$$

and using Eq.(14) one obtains

$$\tilde{M}_{j,l}^{k+1} = A_2 \left(\check{M}_{j-1,l}^k, \check{M}_{j,l}^k, \check{M}_{j+1,l}^k\right) - A_2 \left(\acute{M}_{j-1,l}^k, \acute{M}_{j,l}^k, \acute{M}_{j+1,l}^k\right)$$

which, upon using (J1)-(J4) once more, yields

$$\left|\tilde{M}_{j,l}^{k+1}\right| \leq \omega_{12} \left(\left|\tilde{M}_{j-1,l}^k\right| + \left|\tilde{M}_{j,l}^k\right| + \left|\tilde{M}_{j+1,l}^k\right|\right) \tag{27}$$

for $j = 1, \cdots, N-1$ and $l = 0, \cdots, R$.
Now, define

$$\Lambda^k_{\tilde{C}} = \max_{l=0,\cdots,R} \left|\tilde{C}^k_l\right| \qquad \text{and} \qquad \Lambda^k_{\tilde{M}} = \max_{j=0,\cdots,N\ ;\ l=0,\cdots,R} \left|\tilde{M}^k_{j,l}\right|. \tag{28}$$

Then from Eqs.(24-25)

$$\Lambda^{k+1}_{\tilde{C}} \leq \omega_{13} \left(\Lambda^k_{\tilde{C}} + \Lambda^k_{\tilde{M}}\right) \tag{29}$$

and from Eqs.(26-27)

$$\Lambda^{k+1}_{\tilde{M}} \leq \omega_{14} \ \left(\Lambda^k_{\tilde{C}} + \Lambda^k_{\tilde{M}}\right), \tag{30}$$

which results in

$$\Lambda^k_{\tilde{C}} \leq \omega_{15} \left(\Lambda^0_{\tilde{C}} + \Lambda^0_{\tilde{M}}\right) \tag{31}$$

and

$$\Lambda^k_{\tilde{M}} \leq \omega_{16} \left(\Lambda^0_{\tilde{C}} + \Lambda^0_{\tilde{M}}\right). \tag{32}$$

Eqs.(31-32) easily yield

$$\Lambda^k_{\tilde{C}} = \Lambda^k_{\tilde{M}} = 0,$$

for $k = 0, \cdots, Q$, leading to

$$\tilde{C}^k_l = 0, \tag{33}$$

for $l = 1, \cdots, R$ and $k = 0, \cdots, Q$, as well as

$$\tilde{M}^k_{j,l} = 0 \tag{34}$$

for $j = 0, \cdots, N$, $l = 0, \cdots, R$ and $k = 0, \cdots, Q$. Eqs.(33-34) thus give $\check{C}^k_l = \acute{C}^k_l$ and $\check{M}^k_{j,l} = \acute{M}^k_{j,l}$ which finally results in $\left(\check{C}, \check{M}\right) = \left(\acute{C}, \check{M}\right)$ and concludes proof of uniqueness. Note: ω_1 - ω_{16}, used in the proofs of theorems 1 and 2 of this section are arbitrary positive constants serving as bounds. In the next section, a numerical example is presented which illustrates order of convergence of this finite difference scheme and demonstrates accuracy of the analytically proved results of this section.

4 Numerical Results

For this computational experiment, one uses the following known values for the model parameters:

$$\begin{cases} \lambda = \beta = \rho = x_{\max} = z_{\max} = 1 \\ b\left(C, C_z\right) = C_z, \quad e\left(M, M_x\right) = M_x \\ T_{\max} = \Delta x = \Delta z = 10^{-3}. \end{cases} \tag{35}$$

In order to study order of convergence for the finite difference scheme Eqs.(6,8-10), the following expressions are used for exact solution (C_{ex}, M_{ex}) of Eqs.(1-4):

$$C_{ex}(t,z) = e^{-t}\,(z-1)^2 \tag{36}$$

and

$$M_{ex}(t,x,z) = e^{-t}\,[(x-1)\,(z-1)]^2 \tag{37}$$

which gives the following values for the initial and boundary condition functions:

$$\alpha(z) = (z-1)^2\,, \qquad \gamma(t) = e^{-t}, \qquad \eta(x,z) = [(x-1)(z-1)]^2\,. \tag{38}$$

Further, this value of (C_{ex}, M_{ex}) also leads to the following forcing term for Eq.(1):

$$e^{-t}\left(2z^2 - 2z - 2\right) \tag{39}$$

and the following forcing term for Eq.(2):

$$-2e^{-t}\left(z^2 - 2z + 1\right). \tag{40}$$

Next, the following six distinct values for the time step Δt are chosen:

(I) 8×10^{-8}

(II) 10^{-7}

(III) 1.33×10^{-7}

(IV) 2×10^{-7}

(V) 4×10^{-7}

(VI) 10^{-6},

and for each, the finite difference scheme Eqs.(6,8-10) using parameter values obtained from Eqs.(35-40), is used to obtain the computed model solution denoted by $\left((C_{calc})_l^k, (M_{calc})_{j,l}^k\right)$ from $t = 0$ to $t = T_{\max}$. Finally, for each of these values of Δt, the following quantities are calculated:

$$ABSERC\,(\Delta t) = \max_{l,k} \left| (C_{ex})_l^k - (C_{calc})_l^k \right|$$

and

$$ABSERM\,(\Delta t) = \max_{j,l,k} \left| (M_{ex})_{j,l}^k - (M_{calc})_{j,l}^k \right|.$$

Plotting Δt against $ABSERC\,(\Delta t)$ and $ABSERM\,(\Delta t)$ gives Figures 1 and 2 below, respectively. Both these figures clearly illustrate linear convergence of the finite difference scheme. It may be worthwhile to mention that these computations were executed on a *UNIX* programming environment, with figures plotted using *MATHEMATICA* which proved to be an extremely efficient software for this purpose. This brings us to the conclusion of this section. In the next and final section, which follows after the figures, further research work to be done with this model is discussed.

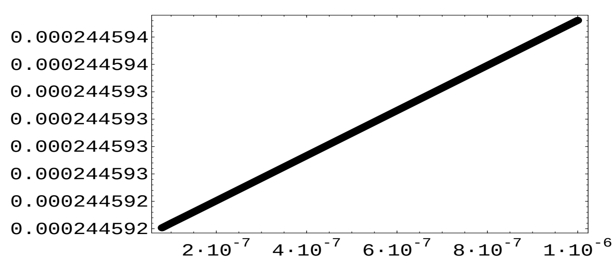


Figure 1: Δt versus $ABSERC$

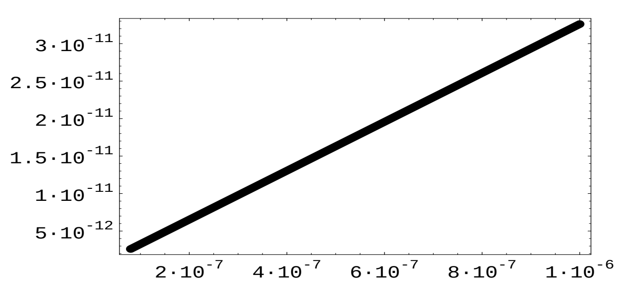


Figure 2: Δt versus $ABSERM$

5 Future Research Work

In the near future an inverse method procedure will be used to numerically estimate infinite-dimensional parameters in this model. This process will involve the minimization of a least-squares cost functional. Convergence results for these numerical parameter approximants will be proved analytically and illustrated graphically. Next, a numerical solution of this model equation will be obtained via a Galerkin approximation method using linear splines, a.k.a. hat functions. Finally, parameter estimation on this model will be performed using this Galerkin method and convergence results for the same will be proved. These results will be submitted for possible publication in refereed journals.

References

[1] A. S. Ackleh and R. R. Ferdinand, A Finite Difference Approximation for a Nonlinear Size-Structured Phytoplankton Dynamics Model, *Quarterly of Applied Mathematics*, **57**, 501-520 (1999).

[2] A. S. Ackleh and R. R. Ferdinand, A Nonlinear Phytoplankton Aggregation Model with Light Shading, *SIAM Journal of Applied Mathematics*, **60**, 316-336 (1999).

[3] R. L. Drake and J. Chen, Contaminant Transport in Parallel Fractured Media: Sudicky and Frind Revisited. Submitted for Publication.

[4] R. R. Ferdinand, Convergence-Uniqueness in a Nonlinear Two-Dimensional Population Diffusion Model, To Appear in the *Journal of Concrete and Applicable Mathematics.*

[5] A. Iserles, *A First Course in the Numerical Analysis of Differential Equations*, Cambridge University Press, Cambridge, UK, 1996.

[6] E. A. Sudicky and E. O. Frind, Contaminant Transport in Fractured Porous Media; Analytic Solutions for a System of Parallel Fractures, *Water Resources Research*, **18(6)**, 1634-1642 (1982).

JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS,VOL.5,NO.2,153-167,2007,COPYRIGHT 2007 EUDOXUS PRESS,LLC

Airline Revenue Optimization Problem: a Multiple Linear Regression Model

Betty Reneé Ferguson
Social Security Administration
6401 Security Blvd.
Baltimore, MD 21235, USA
and
Don Hong
Department of Mathematical Sciences
Middle Tennessee State University
Murfreesboro, TN 37132, USA
dhong@mtsu.edu

Abstract

Airline yield management has been a topic of research since the deregulation of the airline industry in the 1970's. The goal of airline yield management is to optimize seat allocations of a flight among the different fare products. In this paper, we use econometrics modeling to construct market demand functions. Then multiple linear regression is applied to the market demand functions. The use of multiple linear regression allows for an improved discussion of elasticity, cost degradation, and passenger diversion. A model is then constructed to optimize revenue for domestic flights. This paper answers the following specific research questions: How does the use of price elasticities and cross price elasticities improve previous models? Does the use of income elasticity improve the market demand functions? Conditions for optimality are then discussed using the estimated market demand functions.

Key words: airline revenue, cost degradation, econometrics modeling, elasticity, multiple linear regression, optimization.
2000 MSC Codes: 78M50, 62J12.

1 Introduction

Airline yield management, a hot topic of research since the 1970's, is used to optimize seat allocations of a single flight among the different fare products. Most models for airline yield management can be grouped into one of the following two categories: a price discrimination model or a product differentiation model. Price discrimination models assume that when a consumer chooses to purchase a lower priced fare product they do so at no additional cost. If the lower priced fare product requires a purchase of 14 days in advance or any other restrictions applied to a discount purchase, which would not have been encountered by a higher priced fare product, the assumption states that there is no cost to the consumer for accepting more restrictions.

There is an extensive study by Morrison and Winston [6] which estimates the additional costs for accepting more restrictions. Their study supports the need to eliminate the assumption imposed by price discrimination models. The other category, product differentiation models, assumes the demand for fare product i is independent of the demand for fare product j and independent of the price of any other fare products. This paper supports the need to eliminate the assumptions imposed by both the price and product discrimination models.

Botimer and Belobaba [2] introduced a generalized cost model of airline fare product differentiation. Their model for air travel demand in an origin-destination market is extended to include degradation costs and passenger diversion. These extensions eliminate the unrealistic assumptions made by previous price discrimination models and product differentiation models. Degradation costs are the costs to consumers who wish to downgrade to a lower fare product. The lower fare product requires the acceptance of a restriction(s) which may come as a cost to the consumer. Their initial demand function without degradation costs is

$$Q_i = f_i(P_i) - \sum_{j=1}^{i-1} Q_j,$$

where Q_i denotes the number of passengers purchasing fare product i, $f_i(\,\cdot\,)$ is the market demand function for fare product i, P_i is the price of fare product i, and Q_j is the number of passengers held captive to fare products less restricted than fare product i.

Note that fare product $i+1$ is defined to impose more restrictions than fare product i. The model in [2] is designed with the following criteria:
(1) $f_{i+1} < f_i < f_j$ for $j < i$.
(2) The market demand function is a positive function determined by customers, competitors, prices, etc. and exploited by the individual airlines.
(3) The consumers arrive in increasing order of willingness to pay.
(4) Demand for fare product i is derived from unrestricted fare product 1.

To include degradation costs, costs associated with consumers for accepting

more restrictions, their demand function then becomes

$$Q_{i+1} = f_{i+1}(P_{i+1} + \sum_{j=1}^{i-1} c_j) - \sum_{j=1}^{i-1} Q_j,$$

where, c_i is the cost to each consumer for accepting the imposed restrictions. Note that c_1 is zero because there are no restrictions for the full fare product, and thus no cost to consumers. Their model assumes c_i is a constant cost functional form for simplicity reasons. The consumers perceived cost of fare product i is higher than the actual price of fare product i. Thus as Botimer and Belobaba state in [2], "their willingness to purchase fare product i is reduced by c_i as compared to fare product $i-1$. "

The model in [2] for air travel demand is extended to include passenger diversion to eliminate the assumption of previous product differentiation models. This is designed to include a fixed percentage of the expected demand for any fare product. Thus the number of passengers actually purchasing fare product i is represented by q_i,

$$q_i = (1 - \sum_{j=i+1}^{N} d_{ij})Q_i + \sum_{j=1}^{i-1} d_{ji} Q_j,$$

where, d_{ij} denotes the percentage of passengers diverting from fare product i to a more restricted fare product j.

Finally, their optimizing revenue function is constructed only for the linear case of the constant cost model where,

$$P_i = P_0 - a \textstyle\sum_{j=1}^{i-1} Q_j - \sum_{j=1}^{i-1} c_j$$

is strictly nonincreasing. The Lagrange Multiplier Method is used to maximize the following revenue objective function which includes degradation costs and passenger diversion:

$$R = \sum_{i=1}^{N}(1 - \sum_{i=1}^{N} d_{ij})Q_{ij}[P_0 - a\sum_{k=1}^{i} Q_k - \sum_{r=1}^{i} c_r]$$

$$+ \sum_{i=1}^{N} \sum_{j=i+1}^{N} d_{ji} Q_j [P_0 - a \sum_{k=1}^{j} Q_k - \sum_{r=1}^{j} c_r].$$

This model leaves room for improvement as most models do. In [2], the authors mentioned several areas for further research. Under the topic on passenger diversion, the authors suggested the use of cross price elasticity effects in a model. Under the topic on degradation costs, the authors suggested that constant cost formulation does not realistically reflect consumer behavior. That is, "costs incurred may differ by passenger rather than being constant." So there is

a need to determine the effects of passenger behavior. In this paper, comparing to the model in [2] constructed only for the linear case of the constant cost model, we construct a model of market demand function using econometrics modeling that could be used for forecasting consumer behavior in the airline industry. Based on a stratified random sample from the U.S. Department of Transportation's Domestic Airline Fares Consumer Report of 1997, we determine the market demand functions including price and cross price elasticity with and without income elasticity using multiple linear regression. We further analyze these results and determine a generalized objective function. In the final section, we apply Lagrange multiplier method to solve the the airline revenue maximization problems.

2 Market Demand Function

Econometrics modeling of the air travel demand will allow us to observe cost degradation and passenger diversion in action. Instead of fixing the percentage of diversions between the fare products and having a constant degradation cost c_i, we shall model existing behaviors in hopes to be able to better forecast future consumer behavior. For research of econometric modeling, we refer to the book [7]: Econometric Models and Econometric Forecasts.

To understand econometrics modeling and how this can work for airtravel demand, we first recall some basic concepts of elasticity. An elasticity measures the effect on the dependent variable of a 1 percent change in an independent variable. Therefore, we can monitor change of the dependent variable Q_i of a 1 percent change in an independent variable P_i, where Q_i is demand for a product and P_i is the price for this product. This situation is called price elasticity. We can also monitor Q_i of a 1 percent change in another independent variable P_j. This situation is called cross price elasticity. Elasticities are easy to work with due to the facts that their values are unbounded, values may be positive or negative, and are unit-free. A market demand function which includes price elasticity and cross price elasticity may prove to be a more realistic approach to consumer behavior.

Econometric modeling for demand yields the following equation for fare product i

$$Q_i = \beta_{i0} P_i^{\beta_i} \Pi_{i \neq j} P_j^{\beta_{ij}} e^{\epsilon_i},$$

where, Q_i is a continuous, dependent variable representing quantity demanded for fare product i, β_{i0} is an unbounded and unit-free constant, β_i is the price elasticity which is unbounded and unit-free for fare product i, β_{ij} is the unbounded and unit-free cross price elasticity for fare product i by change in price j, P_i is an independent variable representing price of fare product i, and ϵ_i is the error term which assumes a normal distribution.

An econometric model of airline demand shall yield as many equations as there are fare products. Thus for simplicity purposes we shall model demand

aggregating passenger services into three fare products. That is, fare product 1 will have demand function Q_1 representing demand for first class or full fare products, fare product 2 will have demand function Q_2 representing demand for standard economy class or first class with some restrictions, and fare product 3 will have demand function Q_3 representing the demand for discount fareclass or standard economy class with many more restrictions. Q_2 usually requires advance purchase of 3 days while Q_3 usually requires an advance purchase of 14 or more days.

It is clear that we only want to include necessary independent variables. It is not necessary to include P_3 in Q_1 due to the fact that demand in first class is not effected by price changes of fares in the discount fareclass. However, as we will see, we must include P_1, P_2, and P_3 in Q_2 and only P_3 and P_2 in Q_3. Realistically, demand should be effected by the above fareclass and below fareclass changes in price. Thus standard economy fareclass, in the three fare product model, is the only fare product that has two cross price elasticities. In order to linearize the model, we take the natural log of each Q_i yielding the following three fare product market demand functions:

$$\begin{aligned}
\ln Q_1 &= \beta_{10} + \beta_1 \ln P_1 + \beta_{12} \ln P_2 + \epsilon_1, \\
\ln Q_2 &= \beta_{20} + \beta_2 \ln P_2 + \beta_{21} \ln P_1 + \beta_{23} \ln P_3 + \epsilon_2, \\
\ln Q_3 &= \beta_{30} + \beta_3 \ln P_3 + \beta_{32} \ln P_2 + \epsilon_3.
\end{aligned}$$

For the multiple linear regression model, we let $q_i = \ln Q_i$ and $x_i = \ln P_i$. Thus we have the following three market demand functions:

$$\begin{aligned}
q_1 &= \beta_{10} + \beta_1 x_1 + \beta_{12} x_2 + \epsilon_1, \\
q_2 &= \beta_{20} + \beta_2 x_2 + \beta_{21} x_1 + \beta_{23} x_3 + \epsilon_2, \\
q_3 &= \beta_{30} + \beta_3 x_3 + \beta_{32} x_2 + \epsilon_3.
\end{aligned}$$

The assumptions of these multiple linear regressions in [7] are: the relationship between q_i and x_i is linear, the x_i's are non-stochastic variables and in addition, no exact linear relationship exists between two or more independent variables, the error term has zero expected value for all observations, the error term as constant variance for all observations, and the error term is normally distributed.

With these assumptions, we would like to use the model to analyze consumer behavior through the use of price elasticities and cross price elasticities. Therefore, we select a sample of approximately 250 flights to monitor consumer behavior. The data from the 250 flights are the data for q_i's and x_i's in the above model. We use multiple linear regression to search for parameter estimates, β's, that minimize the error sums of squares.

The squared sum of errors is $SSE_i = \sum_j (e_i^{(j)})^2 = \sum_j (q_i^{(j)} - \widehat{q}_i^{(j)})^2$, where $q_i^{(j)}$ is the observed value for the natural log of demand of the flight j and $\widehat{q}_i^{(j)}$

is the predicted value for the natural log of the demand of the flight j. Thus we will have 250 equations for quantity demanded for each of the q_i's with the unknown elasticities β's. Using multiple linear regression we will have one predicted demand function for each of the q_i's:

$$\widehat{q_i} = \widehat{\beta}_{i0} + \widehat{\beta}_i x_i + \widehat{\beta}_{i(i-1)} x_{i-1} + \widehat{\beta}_{i(i+1)} x_{i+1}, \; i = 1, 2, 3.$$

This model needs more than three observations, that is, three or more flights. Multiple linear regression is used to solve for $\widehat{\beta}_i$ and $\widehat{\beta}_{ij}$ for $i \neq j$. Then $\widehat{\beta}_{i0}$ can be solved. The parameter estimates are defined as the following for $\widehat{q}_1$ and similarly for $\widehat{q}_2$ and $\widehat{q}_3$:

$$\widehat{\beta}_1 = \frac{(\sum_{i=1}^{250} x_1^{(i)} q_1^{(i)})(\sum_{i=1}^{250} (x_2^{(i)})^2) - (\sum_{i=1}^{250} x_2^{(i)} q_1^{(i)})(\sum_{i=1}^{250} x_1^{(i)} x_2^{(i)})}{(\sum_{i=1}^{250} (x_1^{(i)})^2)(\sum_{i=1}^{250} (x_2^{(i)})^2) - (\sum_{i=1}^{250} x_1^{(i)} x_2^{(i)})^2},$$

$$\widehat{\beta}_{12} = \frac{(\sum_{i=1}^{250} x_2^{(i)} q_1^{(i)})(\sum_{i=1}^{250} (x_2^{(i)})^2) - (\sum_{i=1}^{250} x_1^{(i)} q_1^{(i)})(\sum_{i=1}^{250} x_1^{(i)} x_2^{(i)})}{(\sum_{i=1}^{250} (x_1^{(i)})^2)(\sum_{i=1}^{250} (x_2^{(i)})^2) - (\sum_{i=1}^{250} x_1^{(i)} x_2^{(i)})^2},$$

and

$$\widehat{\beta}_{10} = \bar{q}_1 - \widehat{\beta}_1 \bar{x}_1 - \widehat{\beta}_{12} \bar{x}_2,$$

where $\bar{q}_1 = \frac{1}{250} \sum_{i=1}^{250} q_1^{(i)}$, and $q_1^{(i)}$ is the observed value for the natural log of demand of the flight i. $\bar{x}_1$ and $\bar{x}_2$ are defined similarly.

3 Main Results and Analysis

To estimate our price and cross price elasticities, a stratified random sample is chosen from U.S. Department of Transportation's Domestic Airline Fares Consumer Report of 1997. This report of the 1,000 largest city-pair markets within the 48 states accounts for approximately 75 percent of all 48-state passengers flights. The 1,000 flights are divided into groups determined by their nonstop distance. A separate simple random sample (SRS) is used to select from the list of flights whose nonstop distance ranges from 100–300 miles, 500–700 miles, 900–1100 miles, 1300–1500 miles, and 1900–2100 miles. The combined SRS in each category yields 250 randomly selected flights; 25 percent of the population of interest.

The prices used for the three fare class model are current prices given from various search engines comprised of www.flyaow.com, www.travelocity.com, and www.bestlodgings.com. These search engines allow us to determine the average prices for each of the three fare classes. The price for standard economy is determined by requiring an advanced purchase of 3 days and 14 days for discount fareclass. The model could also include average median incomes of the city of

departure and the city of arrival. Therefore, in the following, the demand functions will first be solved using price elasticity and cross price elasticity. Then the demand functions will be solved using price, cross price, and income elasticity. Finally, we use statistical analysis to determine whether income elasticity is a useful independent variable for quantity demanded.

Using SAS and applying multiple linear regression to the data for the 250 flights yields the following results:

Market Demand Functions (with price and cross price elasticity)

$$\begin{aligned}\widehat{q_1} &= 9.61 - .253x_1 - .650x_2,\ r^2 = .411,\\ \widehat{q_2} &= 10.8 - .550x_2 - .224x_1 - .123x_3,\ r^2 = .414,\\ \widehat{q_3} &= 10.7 - .172x_3 - .793x_2,\ r^2 = .468,\end{aligned}$$

where r^2 (adj) is .406, .407, and .464 respectively.

Market Demand Functions (with price and cross price elasticity plus income elasticity)

$$\begin{aligned}\widehat{q_1} &= 5.15 - .225x_1 - .669x_2 + .423I,\ r^2 = .415,\\ \widehat{q_2} &= 6.38 - .570x_2 - .216x_1 - .126x_3 + .419I,\ r^2 = .418,\\ \widehat{q_3} &= 6.1 - .169x_3 - .791x_2 + .439I,\ r^2 = .473,\end{aligned}$$

where r^2 (adj) is .408, .409 and .467 respectively and I is the average of the cities (departure and arrival) median family incomes.

Analysis of the model involves several different methods. The methods used here are described in [4]. First, we must analyze the assumptions of the model. One is that the random error term assumes a normal distribution. The histograms of the residuals plotted against each of the independent variables: q_1, q_2, and q_3 indicate a few outliers which may cause a lower than usual measure of fit. Overall, the three histograms appear to have a normal distribution. Thus the analysis of these histograms do not give any indication that the normality assumption of the model has not been met. The normal probability plots of the residuals against each q_1, q_2, and q_3 also show a few possible outliers. According to [4],"an outlier among residuals is one that is far greater than the rest in absolute value and perhaps lies three or four standard deviations or further from the mean of the residuals." Thus there are some concerns from these plots that a few of the flights do not have data that are typical to the rest of the flights. Never the less, the linearity in each of the plots suggests their are no indications that the normality assumption has not been met. Also, the plots of the residuals against the fitted values for q_1, q_2, and q_3 show a few outliers. However, if we remove these outliers, our graph shows constant variance. Thus our assumptions have been met.

Second, we must analyze the cross price elasticity, to verify its importance in the model. We can analyze the analysis of variance tables (ANOVA) to

Table 1: Analysis of Variance: q_1 regress on x_1

Source	DF	Sum Sq.s	Mean Sq.	F-Value	Prob> F
Model	1	52.69197	52.69197	143.321	0.0001
Error	250	91.91282	0.36765		
C Total	251	144.60480			

Root MSE 0.60634 R-square 0.3644 Adj R-sq .03618

Parameter Estimates

Variable	DF	β Estimate	Standard Error
INTERCEP	1	10.006380	0.48941062
X1	1	-0.912193	0.07619605

Variable	T:β =0	Prob > $\|T\|$
INTERCEP	20.446	0.0001
X1	-11.972	0.0001

easily verify the importance of cross price elasticity. The Tables 1–3 are the ANOVA tables for q_1 regressed onto each of the independent variables x_1, x_2, and I. Most obvious from the ANOVA tables are the p-values. The p-value for $q_1 = f(x_1)$ is nearly zero (Table 1) and the p-value for $q_1 = f(x_2)$ is also nearly zero (Table 2). The regression of q_1 on both x_1 and x_2 yields the following sums of squares and the proportion of variations:

$$SSR(x_1) = 48.6686,\ SSR(x_2|x_1) = 10.4507,\ r_1^2 = .3365,\ r_{x_2|x_1}^2 = .1089.$$

These are located in Table 4. It is clear that the demand for the full fare product relies on the price of fare product 2. The data tells us that before x_2 is added to the model, the q_1 with only x_1 in the model had a proportion of variations of .3365. And then, once x_2 is added to the model, the proportion of variations by x_2 after x_1 in the model becomes .1089. It suggests that the model includes cross price elasticity. The necessity for the cross price elasticity can be observed in the same way for q_2 and q_3.

Thus the question still remains, what independent variable is missing from the model? r^2 for all three market demand functions are in the .40 range. For observational data, there are hopes for r^2 to be closer to the .60 range. Therefore, the research was expanded to check for another possible independent variable that may explain the proportion of variations in the quantity demanded. The research expanded to include income elasticity in the model. Since the consumers are purchasing a product, the income for the consumers is an obvious

Table 2: Analysis of Variance: q_1 regress on x_2

Source	DF	Sum Sq.s	Mean Sq.	F-Value	Prob> F
Model	1	58.61769	58.61769	170.426	0.0001
Error	250	85.98710	0.34395		
C Total	251	144.60480			

Root MSE 0.58647 R-square 0.4054 Adj R-sq 0.4030

Parameter Estimates

Variable	DF	β Estimate	Standard Error
INTERCEP	1	9.171135	0.38523367
X2	1	-0.851072	0.06519264

Variable	T:β =0	Prob> $\|T\|$
INTERCEP	23.807	0.0001
X2	-13.055	0.0001

independent variable to analyze. The income for the cities included in the research is the 1999 incomes posted at the website:

verticals.yahoo.com/cities/categories/medfamily.html

which is being used as the best available surrogate.

Using multiple linear regression to include price elasticity, cross price elasticity, and income elasticity we have the three fare product market demand functions listed above. The r^2 and r^2 (adj) as compared to our original three fare product market demand functions did not increase significantly. It is proposed that income elasticity should be dropped from the model and thus not included in the optimality procedure. We can quickly verify the significance, if any, that income may have on q_1 by observing the sums of squares of regression between the three independent variables x_1, x_2, and I.

Now, including the income elasticity we have the following results from the ANOVA tables for q_1:

$$SSR(I|x_1, x_2) = .70009, \quad r^2_{I|x_1,x_2} = .070178$$

Additionally, we have the following variance inflation for the three independent variables and their p-values from the above listed ANOVA tables for q_1:

$$x_1 : 4.4024,\ x_2 : 3.9922,\ I : 1.0172,\ p_{x_1} = .0001,\ p_{x_2} = .0001,\ p_I = .0966.$$

The p-value is the probability for the t-test. The p-value is too high for the independent variable income. The F-partial test also agrees with these results.

Table 3: Analysis of Variance: q_1 regress on I

Source	DF	Sum Sq.s	Mean Sq.	F-Value	Prob> F
Model	1	1.59173	1.59173	2.782	0.0966
Error	250	143.01307	0.57205		
C Total	251	144.60480			

Root MSE 0.75634 R-square 0.0110 Adj R-sq 0.0071

Parameter Estimates

Variable	DF	β Estimate	Standard Error
INTERCEP	1	-2.574539	4.04070019
I	1	.650506	.38997346

Variable	T:β =0	Prob> $\lvert T\rvert$
INTERCEP	-.637	.5246
I	1.668	.0966

Thus income does not make up for the unexplained variation. Observation of the ANOVA tables for q_2 and q_3 can be done in the same way and yields similar results. Thus income elasticity is not a necessary variable for any of the three market demand functions and shall be removed from the model.

The fact still remains that the r^2 for q_1, q_2, and q_3 are .411, .411, and .468 respectively for the sample of 250 flights when I is excluded. Questions arise for the improvement of the model: (1) There might be some terms we should include in the model that can help explain the proportion of variations. For this consideration, an immediate improvement on the model would be to include cross product terms of x_i and x_j in the model of q. This implies considering the family of the exponential models for Q and a flexible functional form would be the translog model (see Final Remarks). (2) Do the unusual observations have such a large effect on the variation? If we exclude a few of the unusual observations, or place less weight on these observations, the variance is nearly constant for each of the three fare product demands. Further analysis of the plot of q_1 versus the predicted value for q_1, if we exclude a few unusual observations then $r^2 = .5023$. Further analysis reveals these unusual observations were data from flights in the northeastern states whose prices for first class and standard economy class were extremely high. Several flights had first class prices above \$1200 and standard economy class above \$900. So further research is needed to improve the model. Research that may involve looking for a key indicator, possibly for the regions of the 48 states since there is some variation in prices

Table 4: Analysis of Variance: q_1 regress on x_1, x_2

Source	DF	Sum Sq.s	Mean Sq.	F-Value	Prob> F
Model	2	59.416	29.708	86.83	0.0001
Error	249	85.189	.342		
C Total	251	144.60480			

Root MSE 0.5849 R-square 0.4141 Adj R-sq 0.406

Parameter Estimates

Variable	DF	β Estimate	Standard Error
INTERCEP	1	9.6117	.4804
X1	1	-.2533	.1658
X2	1	-.6503	.1467

Variable	T:β =0	Prob> $\|T\|$
INTERCEP	20.01	0.0001
X1	-1.53	0.128
X2	-4.43	0.0001

in the different regions.

From the analysis of the market demand functions we have no reason to doubt our normality assumptions and no reason to doubt our optimality model shall exclude income elasticities. These market demand functions reveal consumer behavior within these 250 flights. It is obvious that the changes in price of standard economy class fare products directly effects demand for first class and discount fare class. We can observe the effects of consumer behavior; that is passenger diversion, from these market demand functions. These cross price elasticities offer a clearer picture of the number of passengers who will divert to a lower priced fare product given an increase in price. From the multiple linear regression model, we have observed how consumers react to the degradation costs and the passenger diversion that occurs once we increase or decrease a price of other fare products. Forecasting the future behavior of passenger diversion based on their current behavior is desired.

4 Optimality

The objective now is to maximize revenue, where the decision variables are the prices of the three fare products. Recall, x_i is the natural logarithm of the price of fare product i and q_i is the natural logarithm of the quantity demanded for

fare product i. Initially our objective function defined in terms of price and quantity yields the following problem.

$$\max R = \sum_{i=1}^{3} e^{x_i} e^{q_i},$$

subject to

$$\sum_{i=1}^{3} e^{q_i} \leq \text{capacity}.$$

However, to analyze revenue in terms of price only, we shall rewrite the objective function to include the values of q_i in terms of x_i. Also the constraint shall be rewritten in terms of price. For linearity purposes, the model is designed such that the input for capacity (CAP) will be the logarithm of the capacity of the aircraft. Therefore we have the following problem.

$$\max R = e^{9.61+.747x_1-.650x_2} + e^{10.8+.45x_2-.224x_1-.123x_3} + e^{10.7+.828x_3-.793x_2},$$

subject to:

$$31.11 - .477x_1 - 1.99x_2 - 0.295x_3 - CAP \leq 0.$$

Therefore, we have a nonlinear optimization problem of the constrained case.

Our objective function is clearly convex. However, the objective function is bounded by the capacity of the aircraft. The optimal prices for revenue shall occur when the $\sum_{i=1}^{i=3} e^{q_i}$ is exactly equal to the capacity of the aircraft. Therefore, to find the optimal revenue we apply the Lagrange multiplier method to solve for optimal prices using price elasticity and cross price elasticity as our independent variables.

Our objective function is

$$f(x_1, x_2, x_3) = e^{9.61+.747x_1-.650x_2} + e^{10.8+.45x_2-.224x_1-.123x_3} + e^{10.7+.828x_3-.793x_2}.$$

The constraint is:

$$g(x_1, x_2, x_3) = 31.11 - CAP - .477x_1 - 1.99x_2 - .295x_3 = 0.$$

From the condition $\nabla f = \lambda \nabla g$, We can solve the unknowns x_1, x_2, x_3 and λ, and thus, the optimal prices for revenue.

The data used for the econometric modeling was based on the average daily purchases for the flights. This model is constructed such that an input value for the daily capacity for the aircraft would yield optimal prices for revenue given the market demand functions constructed had a larger r^2. Since the values of x_i is $\ln P_i$, there is a large difference between an $x_1 = 6.4$ and $x_1 = 6.46$, a difference of 38 dollars. Thus when solving the system of equations, we must be very careful to watch the precisions of the digits.

This model appears to be a sound method to use. The model is set up to maximize daily revenues based on price and cross price elasticities. The steps for analysis are clear and the structure of the optimality is clear. Further research into indicator variables for the market demand functions could yield a more accurate model and then using the optimality equations in the same way should prove to output more realistic prices for each of the three fareclasses. The study could be extended to include international and domestic flights. Then the model would have up to N-fare products. The structure would be the same, the market demand functions would be as follows:

$$\widehat{q}_i = \widehat{\beta}_{i0} + \widehat{\beta}_i x_i + \widehat{\beta}_{i(i-1)} x_{i-1} + \widehat{\beta}_{i(i+1)} x_{i+1}.$$

And we would like to

$$\max R = \sum_{i=1}^{N} e^{\widehat{\beta}_{i0} + (1+\widehat{\beta}_i) x_i + \widehat{\beta}_{i(i-1)} x_{i-1} + \widehat{\beta}_{i(i+1)} x_{i+1}},$$

subject to

$$\sum_{j=1}^{N} (\widehat{\beta}_{j0} + \widehat{\beta}_{j1} x_1 + \widehat{\beta}_{j2} x_2 + \cdots + \widehat{\beta}_{jN} x_N) - CAP = 0.$$

5 Final Remarks

1. Despite its simplicity, the linear model is too restrictive and cannot accommodate for the variation in the data. Modern studies of demand and production are usually done in the context of a flexible functional form. Flexible functional forms are used in econometrics because they allow analysts to model second order effects. The most popular flexible functional form is the translog model, which is often interpreted as a second-order approximation to an unknown functional form. Let $\ln y = f(\ln x_1, \cdots, \ln x_k)$. Then its second-order Taylor series around $(x_1, \cdots, x_k) = (1, \cdots, 1)$ is in the form

$$\ln y = \beta_0 + \sum_{i=1}^{k} \beta_k \ln x_i + \frac{1}{2} \sum_{i,j=1}^{k} \gamma_{ij} \ln x_i \ln x_j + \epsilon.$$

Since the value of r^2 in the linear model is at best 0.473, we may consider to apply the translog model for a better fitting. Recently, translog models were used in [3] for the study of productivity change model in the airline industry.

2. Principal component analysis (PCA) involves a mathematical procedure that transforms a number of (possibly) correlated variables into a (smaller) number of uncorrelated variables called principal components. The first principal component accounts for as much of the variability in the data as possible, and

each succeeding component accounts for as much of the remaining variability as possible. In principal component analysis (PCA), the data are fit to a linear model by computing the best linear approximation in the sense of the quadratic error.

Have noted that the r^2 could be improved to a more satisfactory level, we may want to include more variables in the linear model. However, to select as few key independent variables as possible in the model, applying the PCA technique in the modeling will be a good idea. Recently, PCA was applied in [1] for the evaluation of deregulated airline networks with an application to Western Europe.

3. In [5], a new analytical procedure for joint pricing and seat allocation problem was developed using polyhedral graph theoretical approach considering demand forecasts, number of fare classes, and aircraft capacities. Three equivalent models were formulated in the paper: The first model is a 01 integer programming model. The second model is obtained from the first model using the notion of constraint aggregation. The third model is derived by exploiting the special data structure of the first model and utilizing the concepts of split graphs and cutting planes. A decision-support tool was developed for price structure designers to be able to consider a wide variety of possibilities concerning the number of fare classes.

Acknowledgments: This work was supported in part by NSF-IGMS (0408086 and 0552377) and MTSU Research Enhancement Program for Hong.

References

[1] N. Adler and B. Golany, Evaluation of deregulated airline networks using data envelopment analysis combining with principal component analysis with an application to Western Europe, *European Journal of Operational Research*, **132** (2001), 260–273.

[2] T.C. Botimer and P.P. Belobaba, Airline pricing and fare product differentiation: A new theoretical framework, *Journal of the Operational Research Society* **50** (1999), 1085–1097.

[3] R. Ceha and H. Ohta, Productivity change model in the airline industry: A parametric approach, *European Journal of Operation Research* **121** (2000), 641–655.

[4] N.R. Draper and H. Smith, *Applied Regression Analysis*, Third edition, John Wiley and Sons, Inc., New York, 1998.

[5] A. Kuyumcu and A. Garcia-Diaz, A polyhedral graph theory approach to revenue management in the airline industry, *Computers and Industrial Engineering*, **38** (2000), 375–395.

[6] S.R. Morrison and C. Winston, *U.S. Aviation After 20 Years of Deregulation*, Brookings Institution: Washington D.C., 1995.

[7] D.L. Rubinfield and R.S. Pindyck, *Econometric Models and Econometric Forecasts*, McGraw-Hill, Massachusetts, 1998.

Instructions to Contributors
Journal of Concrete and Applicable Mathematics

A quartely international publication of Eudoxus Press, LLC, of TN.

Editor in Chief: George Anastassiou
Department of Mathematical Sciences
University of Memphis
Memphis, TN 38152-3240, U.S.A.

1. Manuscripts hard copies in triplicate, and in English, should be submitted to the Editor-in-Chief:

Prof.George A. Anastassiou
Department of Mathematical Sciences
The University of Memphis
Memphis,TN 38152, USA.
Tel.001. 901.678.3144
e-mail:ganastss@memphis.edu.

Authors may want to recommend an associate editor the most related to the submission to possibly handle it.
Also authors may want to submit a list of six possible referees, to be used in case we cannot find related referees by ourselves.
2. Manuscripts should be typed using any of TEX,LaTEX,AMS-TEX,or AMS-LaTEX and according to EUDOXUS PRESS, LLC. LATEX STYLE FILE. (VISIT www.msci.memphis.edu/~ganastss/jcaam / to save a copy of the style file.)They should be carefully prepared in all respects. Submitted copies should be brightly printed (not dot-matrix), double spaced, in ten point type size, on one side high quality paper 8(1/2)x11 inch. Manuscripts should have generous margins on all sides and should not exceed 24 pages.
3. Submission is a representation that the manuscript has not been published previously in this or any other similar form and is not currently under consideration for publication elsewhere. A statement transferring from the authors(or their employers,if they hold the copyright) to Eudoxus Press, LLC, will be required before the manuscript can be accepted for publication.The Editor-in-Chief will supply the necessary forms for this transfer.Such a written transfer of copyright,which previously was assumed to be implicit in the act of submitting a manuscript,is necessary under the U.S.Copyright Law in order for the publisher to carry through the dissemination of research results and reviews as widely and effective as possible.
4. The paper starts with the title of the article, author's name(s) (no titles or degrees), author's affiliation(s) and e-mail addresses. The affiliation should comprise the department, institution (usually university or company), city, state (and/or nation) and mail code.
The following items, 5 and 6, should be on page no. 1 of the paper.

5. An abstract is to be provided, preferably no longer than 150 words.
6. A list of 5 key words is to be provided directly below the abstract. Key words should express the precise content of the manuscript, as they are used for indexing purposes. The main body of the paper should begin on page no. 1, if possible.
7. All sections should be numbered with Arabic numerals (such as: 1. INTRODUCTION) .
Subsections should be identified with section and subsection numbers (such as 6.1. Second-Value Subheading).
If applicable, an independent single-number system (one for each category) should be used to label all theorems, lemmas, propositions, corrolaries, definitions, remarks, examples, etc. The label (such as Lemma 7) should be typed with paragraph indentation, followed by a period and the lemma itself.
8. Mathematical notation must be typeset. Equations should be numbered consecutively with Arabic numerals in parentheses placed flush right,and should be thusly referred to in the text [such as Eqs.(2) and (5)]. The running title must be placed at the top of even numbered pages and the first author's name, et al., must be placed at the top of the odd numbed pages.
9. Illustrations (photographs, drawings, diagrams, and charts) are to be numbered in one consecutive series of Arabic numerals. The captions for illustrations should be typed double space. All illustrations, charts, tables, etc., must be embedded in the body of the manuscript in proper, final, print position. In particular, manuscript, source, and PDF file version must be at camera ready stage for publication or they cannot be considered.
Tables are to be numbered (with Roman numerals) and referred to by number in the text. Center the title above the table, and type explanatory footnotes (indicated by superscript lowercase letters) below the table.
10. List references alphabetically at the end of the paper and number them consecutively. Each must be cited in the text by the appropriate Arabic numeral in square brackets on the baseline.
References should include (in the following order):
initials of first and middle name, last name of author(s)
title of article,
name of publication, volume number, inclusive pages, and year of publication.
Authors should follow these examples:

<u>Journal Article</u>

1. H.H.Gonska,Degree of simultaneous approximation of bivariate functions by Gordon operators, (journal name in italics) *J. Approx. Theory*, 62,170-191(1990).

<u>Book</u>

2. G.G.Lorentz, (title of book in italics) *Bernstein Polynomials* (2nd ed.), Chelsea,New York,1986.

<u>Contribution to a Book</u>

3. M.K.Khan, Approximation properties of beta operators,in(title of book in italics) *Progress in Approximation Theory* (P.Nevai and A.Pinkus,eds.), Academic Press, New York,1991,pp.483-495.

11. All acknowledgements (including those for a grant and financial support) should occur in one paragraph that directly precedes the References section.
12. Footnotes should be avoided. When their use is absolutely necessary, footnotes should be numbered consecutively using Arabic numerals and should be typed at the bottom of the page to which they refer. Place a line above the footnote, so that it is set

off from the text. Use the appropriate superscript numeral for citation in the text.
13. After each revision is made please again submit three hard copies of the revised manuscript, including in the final one. And after a manuscript has been accepted for publication and with all revisions incorporated, manuscripts, including the TEX/LaTex source file and the PDF file, are to be submitted to the Editor's Office on a personal-computer disk, 3.5 inch size. Label the disk with clearly written identifying information and properly ship, such as:

Your name, title of article, kind of computer used, kind of software and version number, disk format and files names of article, as well as abbreviated journal name.

Package the disk in a disk mailer or protective cardboard. Make sure contents of disks are identical with the ones of final hard copies submitted!

Note: The Editor's Office cannot accept the disk without the accompanying matching hard copies of manuscript. No e-mail final submissions are allowed! The disk submission must be used.

14. Effective 1 Nov. 2005 the journal's page charges are $10.00 per PDF file page. Upon acceptance of the paper an invoice will be sent to the contact author. The fee payment will be due one month from the invoice date. The article will proceed to publication only after the fee is paid. The charges are to be sent, by money order or certified check, in US dollars, payable to Eudoxus Press, LLC, to the address shown in the Scope and
Prices section.

No galleys will be sent and the contact author will receive one(1) electronic copy of the journal issue in which the article appears.

15. This journal will consider for publication only papers that contain proofs for their listed results.

TABLE OF CONTENTS,JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS,VOL.5,NO.2,2007

VOLUME 5,NUMBER 3 JULY 2007

ISSN:1548-5390 PRINT,1559-176X ONLINE

JOURNAL

OF CONCRETE

AND APPLICABLE

MATHEMATICS

EUDOXUS PRESS,LLC

SCOPE AND PRICES OF THE JOURNAL
Journal of Concrete and Applicable Mathematics

A quartely international publication of **Eudoxus Press,LLC**

Editor in Chief: George Anastassiou
Department of Mathematical Sciences,
University of Memphis
Memphis, TN 38152, U.S.A.
ganastss@memphis.edu

The main purpose of the "Journal of Concrete and Applicable Mathematics" is to publish high quality original research articles from all subareas of Non-Pure and/or Applicable Mathematics and its many real life applications, as well connections to other areas of Mathematical Sciences, as long as they are presented in a Concrete way. It welcomes also related research survey articles and book reviews.A sample list of connected mathematical areas with this publication includes and is not restricted to: Applied Analysis, Applied Functional Analysis, Probability theory, Stochastic Processes, Approximation Theory, O.D.E, P.D.E, Wavelet, Neural Networks,Difference Equations, Summability, Fractals, Special Functions, Splines, Asymptotic Analysis, Fractional Analysis, Inequalities, Moment Theory, Numerical Functional Analysis,Tomography, Asymptotic Expansions, Fourier Analysis, Applied Harmonic Analysis, Integral Equations, Signal Analysis, Numerical Analysis, Optimization, Operations Research, Linear Programming, Fuzzyness, Mathematical Finance, Stochastic Analysis, Game Theory, Math.Physics aspects, Applied Real and Complex Analysis, Computational Number Theory, Graph Theory, Combinatorics,
Computer Science Math.related topics,combinations of the above, etc. In general any kind of Concretely presented Mathematics which is Applicable fits to the scope of this journal.
Working Concretely and in Applicable Mathematics has become a main trend in many recent years,so we can understand better and deeper and solve the important problems of our real and scientific world.
"Journal of Concrete and Applicable Mathematics" is a peer- reviewed International Quarterly Journal.
We are calling for papers for possible publication. The contributor should send three copies of the contribution to the editor in-Chief typed in TEX, LATEX double spaced. [See: Instructions to Contributors]

Journal of Concrete and Applicable Mathematics(JCAAM)
ISSN:1548-5390 PRINT, 1559-176X ONLINE.
is published in January,April,July and October of each year by
EUDOXUS PRESS,LLC,
1424 Beaver Trail Drive,Cordova,TN38016,USA,
Tel.001-901-751-3553
anastassioug@yahoo.com
http://www.EudoxusPress.com.
Visit also www.msci.memphis.edu/~ganastss/jcaam.
Webmaster:Ray Clapsadle

Annual Subscription Current Prices:For USA and Canada,Institutional:Print $250,Electronic $220,Print and Electronic $310.Individual:Print $77,Electronic $60,Print

&Electronic $110.For any other part of the world add $25 more to the above prices for Print.
Single article PDF file for individual $8.Single issue in PDF form for individual $25.
The journal carries page charges $10 per page of the pdf file of an article,payable upon acceptance of the article within one month and before publication.
No credit card payments.Only certified check,money order or international check in US dollars are acceptable.
Combination orders of any two from JoCAAA,JCAAM,JAFA receive 25% discount,all three receive 30% discount.

JCAAM is reviewed and abstracted by AMS Mathematical Reviews,MATHSCI,and Zentralblaat MATH.

JCAAM IS A JOURNAL OF RAPID PUBLICATION

BELGIUM
f.bastin@ulg.ac.be
Functional Analysis,Wavelets

5) Paul L.Butzer
RWTH Aachen
Lehrstuhl A fur Mathematik
D-52056 Aachen
Germany
tel.0049/241/80-94627 office,
0049/241/72833 home,
fax 0049/241/80-92212
Butzer@rwth-aachen.de
Approximation Th.,Sampling Th.,Signals,
Semigroups of Operators,Fourier Analysis

6) Yeol Je Cho
Department of Mathematics Education
College of Education
Gyeongsang National University
Chinju 660-701
KOREA
tel.055-751-5673 Office,
055-755-3644 home,
fax 055-751-6117
yjcho@nongae.gsnu.ac.kr
Nonlinear operator Th.,Inequalities,
Geometry of Banach Spaces

7) Sever S.Dragomir
School of Communications and Informatics
Victoria University of Technology
PO Box 14428
Melbourne City M.C
Victoria 8001,Australia
tel 61 3 9688 4437,fax 61 3 9688 4050
sever.dragomir@vu.edu.au,
sever@sci.vu.edu.au
Math.Analysis,Inequalities,Approximation Th.,
Numerical Analysis, Geometry of Banach Spaces,
Information Th. and Coding

8) Angelo Favini
Universita di Bologna
Dipartimento di Matematica
Piazza di Porta San Donato 5
40126 Bologna, ITALY
tel.++39 051 2094451
fax.++39 051 2094490
favini@dm.unibo.it
Partial Differential Equations, Control Theory,
Differential Equations in Banach Spaces

tel.805-893-4869
rachev@pstat.ucsb.edu
AND
Chair of Econometrics and Statistics
School of Economics and Business Engineering
University of Karlsruhe
Kollegium am Schloss,Bau II,20.12,R210
Postfach 6980,D-76128,Karlsruhe,Germany
tel.011-49-721-608-7535
rachev@lsoe.uni-karlsruhe.de
Mathematical and Empirical Finance,
Applied Probability, Statistics and Econometrics

25) John Michael Rassias
University of Athens
Pedagogical Department
Section of Mathematics and Infomatics
20, Hippocratous Str., Athens, 106 80, Greece

Address for Correspondence
4, Agamemnonos Str.
Aghia Paraskevi, Athens, Attikis 15342 Greece
jrassias@primedu.uoa.gr
jrassias@tellas.gr
Approximation Theory,Functional Equations,
Inequalities, PDE

26) Paolo Emilio Ricci
Universita' degli Studi di Roma "La Sapienza"
Dipartimento di Matematica-Istituto "G.Castelnuovo"
P.le A.Moro,2-00185 Roma,ITALY
tel.++39 0649913201,fax ++39 0644701007
riccip@uniroma1.it,Paoloemilio.Ricci@uniroma1.it
Orthogonal Polynomials and Special functions,
Numerical Analysis, Transforms,Operational Calculus,
Differential and Difference equations

27) Cecil C.Rousseau
Department of Mathematical Sciences
The University of Memphis
Memphis,TN 38152,USA
tel.901-678-2490,fax 901-678-2480
ccrousse@memphis.edu
Combinatorics,Graph Th.,
Asymptotic Approximations,
Applications to Physics

28) Tomasz Rychlik
Institute of Mathematics
Polish Academy of Sciences
Chopina 12,87100 Torun, Poland
T.Rychlik@impan.gov.pl
Mathematical Statistics,Probabilistic Inequalities

9) A.M.Fink
Department of Mathematics
Iowa State University
Ames,IA 50011-0001,USA
tel.515-294-8150
fink@math.iastate.edu
Inequalities,Ordinary Differential
Equations

10) Sorin Gal
Department of Mathematics
University of Oradea
Str.Armatei Romane 5
3700 Oradea,Romania
galso@uoradea.ro
Approximation Th.,Fuzzyness,Complex
Analysis

11) Jerome A.Goldstein
Department of Mathematical Sciences
The University of Memphis,
Memphis,TN 38152,USA
tel.901-678-2484
jgoldste@memphis.edu
Partial Differential Equations,
Semigroups of Operators

12) Heiner H.Gonska
Department of Mathematics
University of Duisburg
Duisburg,D-47048
Germany
tel.0049-203-379-3542 office
gonska@informatik.uni-duisburg.de
Approximation Th.,Computer Aided
Geometric Design

13) Dmitry Khavinson
Department of Mathematical Sciences
University of Arkansas
Fayetteville,AR 72701,USA
tel.(479)575-6331,fax(479)575-8630
dmitry@uark.edu
Potential Th.,Complex Analysis,Holomorphic
PDE,
Approximation Th.,Function Th.

14) Virginia S.Kiryakova
Institute of Mathematics and Informatics
Bulgarian Academy of Sciences
Sofia 1090,Bulgaria
virginia@diogenes.bg
Special Functions,Integral Transforms,
Fractional Calculus

15) Hans-Bernd Knoop

29) Bl. Sendov
Institute of Mathematics and Informatics
Bulgarian Academy of Sciences
Sofia 1090,Bulgaria
bsendov@bas.bg
Approximation Th.,Geometry of Polynomials,
Image Compression

30) Igor Shevchuk
Faculty of Mathematics and Mechanics
National Taras Shevchenko
University of Kyiv
252017 Kyiv
UKRAINE
shevchuk@univ.kiev.ua
Approximation Theory

31) H.M.Srivastava
Department of Mathematics and Statistics
University of Victoria
Victoria,British Columbia V8W 3P4
Canada
tel.250-721-7455 office,250-477-6960 home,
fax 250-721-8962
harimsri@math.uvic.ca
Real and Complex Analysis,Fractional Calculus
and Appl.,
Integral Equations and Transforms,Higher
Transcendental
Functions and Appl.,q-Series and q-Polynomials,
Analytic Number Th.

32) Ferenc Szidarovszky
Dept.Systems and Industrial Engineering
The University of Arizona
Engineering Building,111
PO.Box 210020
Tucson,AZ 85721-0020,USA
szidar@sie.arizona.edu
Numerical Methods,Game Th.,Dynamic Systems,
Multicriteria Decision making,
Conflict Resolution,Applications
in Economics and Natural Resources
Management

33) Gancho Tachev
Dept.of Mathematics
Univ.of Architecture,Civil Eng. and Geodesy
1 Hr.Smirnenski blvd
BG-1421 Sofia,Bulgaria
gtt_fte@uacg.bg
Approximation Theory

34) Manfred Tasche
Department of Mathematics

Institute of Mathematics
Gerhard Mercator University
D-47048 Duisburg
Germany
tel.0049-203-379-2676
knoop@math.uni-duisburg.de
Approximation Theory,Interpolation

16) Jerry Koliha
Dept. of Mathematics & Statistics
University of Melbourne
VIC 3010,Melbourne
Australia
koliha@unimelb.edu.au
Inequalities,Operator Theory,
Matrix Analysis,Generalized Inverses

17) Mustafa Kulenovic
Department of Mathematics
University of Rhode Island
Kingston,RI 02881,USA
kulenm@math.uri.edu
Differential and Difference Equations

18) Gerassimos Ladas
Department of Mathematics
University of Rhode Island
Kingston,RI 02881,USA
gladas@math.uri.edu
Differential and Difference Equations

19) Rupert Lasser
Institut fur Biomathematik & Biomertie,GSF
-National Research Center for environment
and health
Ingolstaedter landstr.1
D-85764 Neuherberg,Germany
lasser@gsf.de
Orthogonal Polynomials,Fourier Analysis,
Mathematical Biology

University of Rostock
D-18051 Rostock
Germany
manfred.tasche@mathematik.uni-rostock.de
Approximation Th.,Wavelet,Fourier Analysis,
Numerical Methods,Signal Processing,
Image Processing,Harmonic Analysis

35) Chris P.Tsokos
Department of Mathematics
University of South Florida
4202 E.Fowler Ave.,PHY 114
Tampa,FL 33620-5700,USA
profcpt@math.usf.edu,profcpt@chuma1.cas.usf.edu
Stochastic Systems,Biomathematics,
Environmental Systems,Reliability Th.

36) Lutz Volkmann
Lehrstuhl II fuer Mathematik
RWTH-Aachen
Templergraben 55
D-52062 Aachen
Germany
volkm@math2.rwth-aachen.de
Complex Analysis,Combinatorics,Graph Theory

JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS,VOL.5,NO.3,181-196,2007,COPYRIGHT 2007 EUDOXUS PRESS,LLC

On the numerical approximation to π

George W. Grossman
Department of Mathematics
Central Michigan University
Mount Pleasant, MI
48859

gross1gw@cmich.edu
Subject class: Primary: 01A99, Secondary 65Q05
Keywords: History of mathematics, Recurrence relations

Abstract. In the present paper we consider recurrence equations that can be used to approximate π. Analysis of each approximation shows that overall accuracy can be improved by taking a linear combination of the results of the two equations, after finitely many steps. Additionally, one of the recurrence relations is derived using the angle-bisector theorem. The analysis is based on Taylor series applied to trigonometric functions. There are also several new results presented relating cubic equations and trigonometric functions of half-angles. We determine a bound on the 'closeness' of trigonometric differences for small angles using a Lipschitz condition and give examples. We employ MAPLE software in the present work to facilitate computation.

1. Introduction

The present paper is related to historical developments in mathematics in that it shows the use of the Taylor series to improve the accuracy of a classical approximation to π and also to analyze area of regular polygons. An application of Taylor polynomials to Newton's integral [3],

$$\pi = \frac{3\sqrt{3}}{4} + 24\int_0^{0.25} \sqrt{x - x^2}dx \ , \tag{1.1}$$

can be shown by expanding (1.1) about $x = 1/4$ so that,

$$\pi = \frac{3\sqrt{3}}{4} + 24\int_0^{0.25} \frac{1}{4}\sqrt{3} + \frac{1}{3}\sqrt{3}\left(x - \frac{1}{4}\right) - \frac{8}{9}\sqrt{3}\left(x - \frac{1}{4}\right)^2 + \ldots dx \ . \tag{1.2}$$

Each successive term in the expansion of $\sqrt{x - x^2}$ gives a new approximation to π. The integral curve

$$Y = \int \sqrt{1 - x^2}dx = \frac{1}{2}\sin^{-1} x + \frac{x}{2}\sqrt{1 - x^2} \ ,$$

has ordinate $Y(1) = \pi/4$. In the late 19th century, an instrument called the *integraph* [15], and invented by Abdank-Abakanowicz, could be used to trace such a curve and so construct approximations to π.

The π approximation problem was considered by Archimedes's [10], [14], [18] (c. 250 B.C.) and still has significance almost two thousand three hundred years later. Archimedes approximated π by finding the perimeters of regular polygons circumscribing and inscribing the unit circle. The polygons in the paper are denoted $p_n, \ P_n$, respectively, each having 2^{n+1} sides.

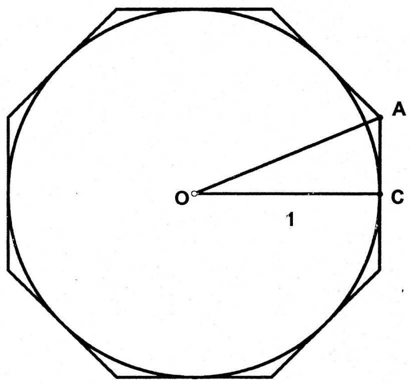


Figure 1: P_2 , regular polygon of 8 sides.

(Fig. 1 shows P_2, fig. 6 shows P_1 and p_1.) We observe that P_n, respectively p_n can be subdivided into 2^{n+1} congruent, isosceles triangles each having interior angles $\pi/2^n$ at the vertex (origin) and the other two equal angles at the circumference are equal to (in radians),

$$\frac{\pi}{2}\left(1-\frac{1}{2^n}\right) .$$

It is noted that the interior angle is halved for successive values of n. Accuracy in approximating π was limited [5], [18] and Archimedes obtained the following approximation,

$$3\frac{10}{71} < \pi < \frac{22}{7}.$$

Archimedes's recurrence relation [1] approximates the perimeter of the unit circle, employs the harmonic-geometric means with initial guess $a_1,\ b_1$, and is given by

$$a_{2n} = \frac{2a_n\, b_n}{a_n + b_n}, \tag{1.3}$$

$$b_{2n} = \sqrt{a_{2n}\, b_n}\ ,\ 2n = 2, 4, 8, \ldots\ , \tag{1.4}$$

such that

$$\lim_{n\to\infty} a_{2n} = 2\pi\ ,\ \lim_{n\to\infty} b_{2n} = 2\pi\ .$$

The advent of electronic calculators and computers allows programming of such approximation methods, including the Brent-Salamin [6] formulas, to obtain (quadratically convergent) approximations to π given by,

$$\begin{aligned} \pi &= \frac{4[M(1,1/\sqrt{2})]^2}{1-\sum_{j=1}^{\infty} 2^{j+1} d_j}, \\ &= \frac{4[M(1,1/\sqrt{2})]^2}{1-\sum_{j=1}^{\infty} 2^{j+1} c_j^2}\ , \end{aligned}$$

where formulas for arithmetic-geometric mean, denoted $M(a_0, b_0) = \lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n$, with initial guess $a_0,\ b_0$, are given by

$$\begin{aligned} a_{n+1} &= \frac{1}{2}(a_n + b_n), \\ b_{n+1} &= \sqrt{a_n\, b_n}, \end{aligned}$$

such that

$$\begin{aligned} c_{n+1} &= \frac{1}{2}(a_n - b_n), \\ d_n &= a_n^2 - b_n^2. \end{aligned}$$

We note that $M(a,b) = M(b,a)$ so that

$$M(1, 1/\sqrt{2}) = \frac{1}{\sqrt{2}} M(\sqrt{2}, 1) = \frac{1}{\sqrt{2}} M(1, \sqrt{2})$$

and from [19]

$$M(1, \sqrt{2})^{-1} = \frac{2}{\pi} \int_0^1 \frac{dt}{\sqrt{1-t^4}} \ .$$

Srinivasa Ramanujan [13] employed modular functions [9] to generate approximations to π, resulting in infinite series,

$$\frac{4}{\pi} = \frac{3}{2} - \frac{23}{2^3}\frac{1}{2}\frac{1\cdot 3}{4^2} + \frac{43}{2^5}\frac{1}{2}\frac{3}{4}\frac{1\cdot 3\cdot 5\cdot 7}{4^2\cdot 8^2} - \cdots \ .$$

The Bailey-Borwein-Plouffe algorithm (BBP) [7], another highly accurate formula for calculating π, and so-called 'digit extraction method', is given by

$$\pi = \sum_{n=0}^{\infty} \frac{1}{16^n} \left[\frac{4}{8n+1} - \frac{2}{8n+4} - \frac{1}{8n+5} - \frac{1}{8n+6} \right].$$

Half-angle (fig. 5 illustrates this idea) formulas were also employed by Viete [11] who derived the following infinite sequence,

$$\frac{2}{\pi} = \lim_{n\to\infty} (u_1 u_2 u_3 \cdots u_n) = \sqrt{\frac{1}{2}} \sqrt{\frac{1}{2} + \frac{1}{2}\sqrt{\frac{1}{2}}} \sqrt{\frac{1}{2} + \frac{1}{2}\sqrt{\frac{1}{2} + \frac{1}{2}\sqrt{\frac{1}{2}}} \cdots } \ , \tag{1.5}$$

such that,

$$u_1 = \cos\frac{\pi}{4} = \sqrt{\frac{1}{2}}, \quad u_n = \cos\frac{\pi}{2^{n+1}} = \sqrt{\frac{1}{2}(1 + u_{n-1})} \ , \ n > 1.$$

This is a case of a more general formula due to Euler [11],

$$\frac{\sin\theta}{\theta} = \lim_{n\to\infty} (u_1 u_2 u_3 \cdots u_n) \ ,$$

where

$$u_n = \cos\frac{\theta}{2^n} \ , \ n \geq 1.$$

We note that (1.5) is obtained by setting $\theta = \pi/2$ in u_n above.

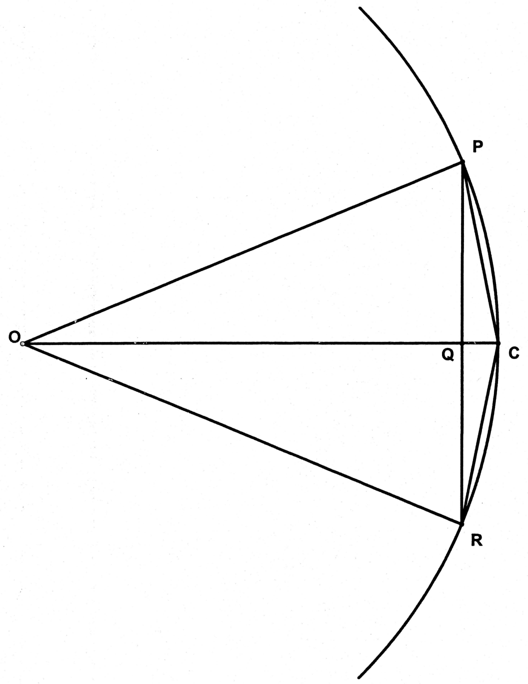


Figure 2: $|PC| = S_{n+1}$, $|OQ| = h_n$, $|OC| = 1$, $|PR| = S_n$

Convexity and symmetry [12] are properties of regular polygons. It is known [17] (sect 3.) because of convexity the approximation to π by $area(P_n)$ is more accurate than by $area(p_n)$. An exercise in [17] yields

$$area(p_n) < \pi < area(p_{n+1}) + \frac{\pi}{2^n} \ . \tag{1.6}$$

A similar error bound is also explored in sect. 6.

The paper is organized as follows. in sect. 2 we discuss recurrence equations. In sect. 3 we analyze the convergence of Archimedes' recurrence relation. In sect. 4 area formulas related to triangular subregions of the regular polygons p_n, P_n are determined and, several inequalities are derived. In sect. 5 algebraic relations are derived related to the congruent triangular subregions of p_n, P_n. In sect. 6 differences of trigonometric functions for small angles are determined, a numerical algorithm is given and accuracy is analyzed. Several examples are presented.

2. Recurrence equations

A recurrence equation [17] is given for the side-length denoted $s_n, \ n > 1, \ s_1 = \sqrt{2}$, of a regular polygon p_n inscribing the unit disk,

$$s_{n+1} = \sqrt{2 - \sqrt{4 - s_n^2}} \ . \tag{2.1}$$

It can easily be checked that $s_n = 2\sin(\pi/2^{n+1})$ $, n \geq 1$ satisfies the recursion (2.1) which can be derived from two applications of Pythagorean identities relating the altitude h_n and side-lengths $s_n, \ s_{n+1}$ of the 2^{n+1}, (2^{n+2}) congruent isosceles triangles of $p_n, \ (p_{n+1})$, i.e., (fig. 2)

$$(s_{n+1})^2 = (1 - h_n)^2 + \left(\frac{s_n}{2}\right)^2 , \ h_n^2 + \left(\frac{s_n}{2}\right)^2 = 1 \ . \tag{2.2}$$

Elimination of h_n from (2.2) gives (2.1) using elementary geometry. A natural question arises as to the form of the recursion for the side-length denoted t_n of the circumscribed polygon P_n and this is known also [11] and written in terms of the tangent function $t_n = 2\tan(\pi/2^{n+1})$, $n \geq 1$ such that

$$t_{n+1} = \frac{t_n}{1+\sqrt{1+(t_n/2)^2}} \ . \tag{2.3}$$

t_n and s_n, respectively, are related to the half-angle formulas,

$$\tan\theta/2 = \frac{\sin\theta}{1+\cos\theta} \ , \ \sin\frac{\theta}{2} = \frac{1}{\sqrt{2}}\sqrt{1-\cos\theta} \ , \tag{2.4}$$

by observing that

$$\tan\theta/2 = \frac{\sin\theta}{1+\cos\theta} = \frac{\tan\theta}{1+\sqrt{1+(\tan\theta)^2}} \ , \tag{2.5}$$

and

$$2\sin\frac{\theta}{2} = \sqrt{2}\sqrt{1-\cos\theta} = \sqrt{2-\sqrt{4-(2\sin\theta)^2}} \ . \tag{2.6}$$

The half-length defined as $b_n = t_n/2$ satisfies

$$b_{n+1} = \frac{b_n}{1+\sqrt{1+b_n^2}} \ . \tag{2.7}$$

We note that (2.7) has an interesting derivation that employs a recursive idea and the angle-bisector theorem, (in sects. 5, 6.) We note that (2.1) can be used to approximate π, that is, $area(unitdisk) = \pi \approx 2^n s_n$. In the case of (2.7) we have $area(unitdisk) = \pi \approx 2^n t_n$.

3. Harmonic and Geometric Means

In this section we derive convergence properties (see [11]) for a different derivation) of Archimedes' recurrence relation. It can be shown [1] that s_n and t_n defined in (2.1, 2.3) (replacing b_n and a_n respectively) satisfy (1.3, 1.4). Employing $a_{2n} = b_{2n}^2/b_n$ from (1.4) in (1.3) yields,

$$b_{2n} = \frac{\sqrt{2}b_n^{3/2}}{\sqrt{b_n+b_{n/2}}}, \tag{3.1}$$

$$a_{2n} = \frac{2b_n^2}{b_n+b_{n/2}} = \frac{2}{b_{n/2}/b_n+1}b_n. \tag{3.2}$$

By solving for b_{2n}/b_n in (3.1) and expanding about 1, we obtain

$$\frac{b_{2n}}{b_n} = \frac{\sqrt{2}}{\sqrt{1+b_{n/2}/b_n}} = 1 - 0.25\left(\frac{b_{n/2}}{b_n}-1\right) + O\left(\frac{b_{n/2}}{b_n}-1\right)^2 . \tag{3.3}$$

From (3.3) we obtain,

$$\left|\left(\frac{b_{2n}}{b_n}-1\right) \div \left(\frac{b_{n/2}}{b_n}-1\right)\right| \approx \frac{1}{4} \ , \tag{3.4}$$

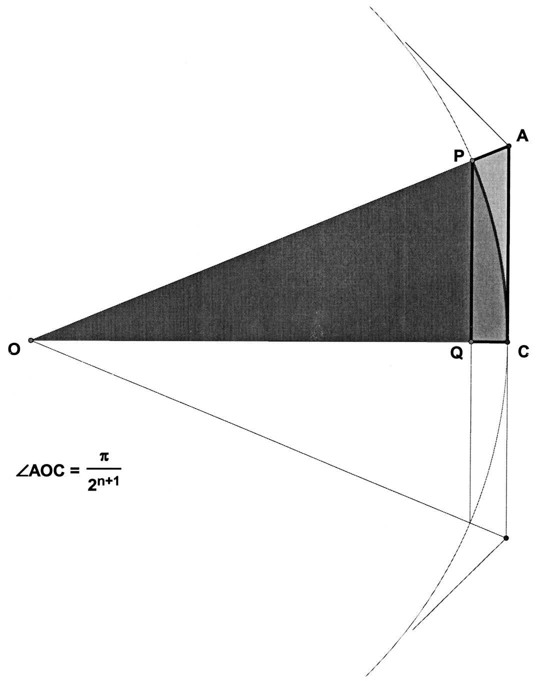


Figure 3: $\triangle OAC$, $\triangle OPQ$, APC, PQC, APQC.

so that,

$$\left|\frac{b_{2n} - b_n}{b_n - b_{n/2}}\right| \approx \frac{1}{4}. \tag{3.5}$$

From (3.5), the convergence of the subsequence $b_1, b_2, b_4, \ldots, b_{2^n}, \ldots$ is linear and also from (3.2), the sequence $a_1, a_2, a_4, \ldots, a_{2^n}, \ldots$ converges linearly. Thus, the sequences $\{s_n\}$ and $\{t_n\}$ converge linearly. In sect 6 it is shown that the number of significant digits in the approximation to π by $2^n s_n$ and $2^n t_n$ is proportional to n.

We note that convergence of an infinite sequence $\{x_k\}$ is said to be linear [16] if

$$\frac{\triangle x_k}{\triangle x_{k-1}} \approx const = C_L \ , \ 0 < C_L < 1 \ , \triangle x_k = x_k - x_{k-1}.$$

In the next section we obtain elementary results concerning the areas of various subregions (fig. 3) of the regular polygons p_n, P_n and make several conclusions.

4. Area formulas

From fig. 1 we see that $area(P_2) = 2^4 area(\triangle AOC)$ and $\angle AOC = \pi/2^3$. Moreover, it follows that,

$$\pi = \lim_{n\to\infty} area(P_n) = \lim_{n\to\infty} 2^{n+2} area(\triangle OAC),$$

such that $\angle AOC = \pi/2^{n+1}$. Similar results apply for p_n and $\triangle OPQ$. The following results refer to fig. 3 and curvilinear (without $\triangle$ notation) or triangular regions within p_n, P_n so that in the foregoing it is assumed for convenience that these regions depend implicitly on n and p_n or P_n. For $n \geq 1$ we have the following two lemmas,

Lemma 4.1.

$$\begin{aligned}
(4.1)\qquad area(\triangle OPQ) &= \frac{1}{2}\sin\frac{\pi}{2^{n+1}}\cos\frac{\pi}{2^{n+1}} = \frac{1}{4}\sin\frac{\pi}{2^n} = \frac{1}{8}s_{n-1}\ ,\ (s_0 = 2)\ , \\
(4.2)\qquad area(\triangle OAC) &= \frac{1}{2}\tan\frac{\pi}{2^{n+1}} = \frac{1}{4}t_n\ , \\
(4.3)\qquad area(APQC) &= \frac{1}{2}\frac{\sin^3(\pi/2^{n+1})}{\cos(\pi/2^{n+1})} = \frac{1}{2}\sin^2\frac{\pi}{2^{n+1}}\tan\frac{\pi}{2^{n+1}} = \left(\frac{s_n}{4}\right)^2 t_n\ , \\
(4.4)\qquad area(PQC) &= \frac{\pi}{2^{n+2}} - \frac{1}{4}\sin\frac{\pi}{2^n} = \frac{\pi}{2^{n+2}} - \frac{1}{8}s_{n-1}\ , \\
(4.5)\qquad area(APC) &= \frac{1}{2}\sin^2\frac{\pi}{2^{n+1}}\tan\frac{\pi}{2^{n+1}} - \frac{\pi}{2^{n+2}} + \frac{1}{4}\sin\frac{\pi}{2^n} = \frac{1}{2}\tan\frac{\pi}{2^{n+1}} - \frac{\pi}{2^{n+2}} \\
&= \frac{1}{4}t_n - \frac{\pi}{2^{n+2}}.
\end{aligned}$$

Proof. From fig. 3, $|OP| = |OC| = 1$, $\angle AOC = \pi/2^{n+1}$, and trigonometric formulas for the right triangle yields (4.1, 4.2). The difference of (4.1, 4.2) yields (4.3). To obtain (4.4) we can evaluate

$$\int_{\cos\theta}^{1}\sqrt{1-x^2}dx\ ,\ \theta = \frac{\pi}{2^{n+1}}.$$

To obtain (4.5) we can evaluate

$$\int_{\cos\theta}^{1}\tan\theta\ x - \sqrt{1-x^2}dx\ ,\ \theta = \frac{\pi}{2^{n+1}}.$$

□

It can be shown by Taylor series and (4.4, 4.5) that

$$(4.6)\qquad area(APC) = area(PQC) - \left(\frac{\pi}{2^{n+1}}\right)^3 + \frac{2}{15}\left(\frac{\pi}{2^{n+1}}\right)^5 + O\left(\frac{\pi}{2^{n+1}}\right)^7\ .$$

We also have,

Lemma 4.2.

$$(4.7)\qquad area(APC) < area(PQC)\ , n \geq 2\ ,$$

$$(4.8)\qquad 2^{n+1}area(\triangle OAC) - \pi < \pi - 2^{n+1}area(\triangle PQO)\ , n \geq 2\ .$$

Proof. We observe that (4.6) suggests (4.7) for large enough n, but this requires looking at the Taylor series remainder more closely to find the minimum value of n, however we show that the areas defined by $A_1 = area(segment(PC)), A_2 = area(\triangle APR)$ satisfy $A_1 > A_2$ for $n \geq 2$ so that (figs. 3, 7)

$$\begin{aligned}
area(PQC) &= A_1 + area(\triangle PQC) = A_1 + area(\triangle PCR) \\
&> area(PCR) + A_2 = area(PCR) + area(\triangle APR) = area(APC).
\end{aligned}$$

It follows that for $\cos\theta \le x \le 1,\ \theta = \pi/2^{n+1}$,

$$\sin\theta \ \ge\ \sqrt{1-x^2} \ \ge\ \frac{\sin\theta}{\cos\theta - 1}(x-1)\ .$$

These inequalities (with strict inequality in the open interval $(\cos\theta, 1)$) give $area(\triangle PCR) > area(PCR)$. Also, the arc of the unit circle over the interval $[0,\theta]$ lies above the straight line segment PC, (fig. 7). We have for $n \ge 2$,

$$\begin{aligned}
A_1 &= \int_{\cos\theta}^{1} \sqrt{1-x^2} - \frac{\sin\theta}{\cos\theta - 1}(x-1)dx \\
&= \frac{\pi}{2^{n+2}} - \frac{1}{2}\sin\left(\frac{\pi}{2^{n+1}}\right) \\
&= \frac{1}{2}\left[\left(\frac{\pi}{2^{n+1}}\right)^3 \frac{1}{3!} - \left(\frac{\pi}{2^{n+1}}\right)^5 \frac{1}{5!} + \dots\right] \\
&> \left(\frac{\pi}{2^{n+1}}\right)^3 \frac{1}{12}\left(1 - \left(\frac{\pi}{2^{n+1}}\right)^2 \frac{1}{20}\right) \\
&> \frac{1}{8}\tan\left(\frac{\pi}{2^{n+1}}\right)\left(\frac{\pi}{2^{n+1}}\right)^4 \\
&> \frac{1}{2}\tan\left(\frac{\pi}{2^{n+1}}\right)\left[\left(\frac{\pi}{2^{n+1}}\right)^2 \frac{1}{2!} - \left(\frac{\pi}{2^{n+1}}\right)^4 \frac{1}{4!} + \dots\right]^2 \\
&= \frac{1}{2}\tan\left(\frac{\pi}{2^{n+1}}\right)\left(1 - \cos\left(\frac{\pi}{2^{n+1}}\right)\right)^2 \\
&= \frac{1}{2}(\tan\theta - \sin\theta)(1-\cos\theta) = A_2\ .
\end{aligned}$$

We rewrite (4.8) as $2^{n+1}[area(\triangle OAC) + area(\triangle PQO)] < 2\pi$, which follows from $2^{n+1}[area(APC) + area(\triangle PQO)] < \pi$ by(4.7) and $2^{n+1}[area(PQC) + area(\triangle PQO)] = \pi$. ☐

For $n = 1$, $A_1 < A_2$. As noted earlier an outer approximation to π by $area(P_n)$, is a more accurate approximation than an inner approximation by $area(p_n)$. In the next section we employ the law of sines that yields several algebraic identities.

5. Algebraic relations

Referring to fig. 4 we find that $|OC| = 1$, $\angle BOC = \angle AOC$,$\angle ACO = \pi/2$,$a + b = |AC|$,$|OA| = c$,$|OB| = d$. We state a version of the well-known angle-bisector theorem [2] and give an alternate proof by the law of sines, (fig. 4).

Lemma 5.1. *Consider $\triangle AOC$, $\triangle AOB$ and $\triangle BOC$ such that $\angle AOC$ is bisected into $\angle AOB$ and $\angle BOC$. Denote $a = |AB|$,$b = |BC|$. We have*

$$\frac{a}{b} = c\ ,b \neq 0. \tag{5.1}$$

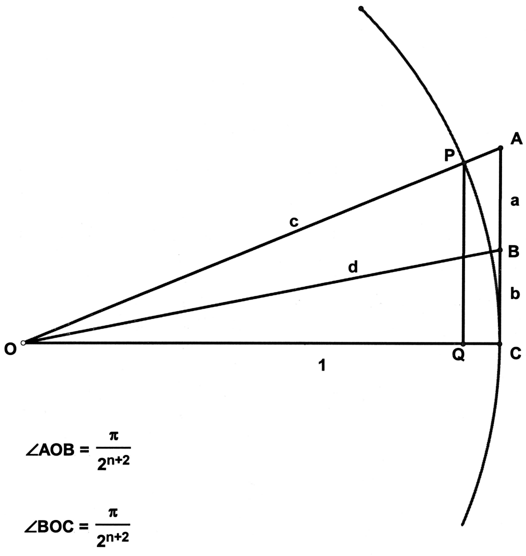


Figure 4: C = |OA| , a = |AB| , b = |BC| , d = |OB| , |OP| = |OC| = 1

Proof. From the law of sines we obtain

$$\frac{b}{\sin\theta/2} = \frac{d}{\sin\pi/2} = \frac{1}{\sin(\pi-\theta)/2} = \frac{1}{\cos\theta/2}\ , \tag{5.2}$$

$$\frac{c}{\sin(\pi+\theta)/2} = \frac{a}{\sin\theta/2} = \frac{d}{\sin(\pi/2-\theta)} = \frac{d}{\cos\theta}\ . \tag{5.3}$$

The result (5.1) follows by choosing the ratio of terms involving $a, b, c, 1$ in (5.2-5.3). □

We remark that in [2] the proof of the angle-bisector theorem uses elementary geometry. Fig. 5 shows P_1 and the manner in which angles (beginning with 90 degrees) are bisected repeatedly with respect to P_1. (Fig. 6 shows P_1, p_1.)

Employing the Pythagorean identity and (5.1) we obtain

$$c^2 = 1 + (a+b)^2 = (a/b)^2. \tag{5.4}$$

For completeness, we find using (5.1, 5.4), the relations between $a,\ b,\ c,\ d$ given in the following.

From (5.4) we obtain $a + b = (a-b)/b^2$ which yields with (5.1) and elementary algebra,

$$a \;=\; \frac{b(1+b^2)}{1-b^2} = c\sqrt{\frac{c-1}{c+1}} = \frac{\sqrt{(d^2-1)}d^2}{2-d^2}\ ,0 < b < 1\ ,1 < d < c < \sqrt{2}\ , \tag{5.5}$$

$$c \;=\; \frac{1+b^2}{1-b^2} = \frac{1}{2/d^2-1}\ ,0 < b < 1\ ,1 < d < \sqrt{2}, \tag{5.6}$$

$$b \;=\; \sqrt{\frac{c-1}{c+1}} = \sqrt{d^2-1}\ ,1 < d < c < \sqrt{2}, \tag{5.7}$$

$$d \;=\; \sqrt{1+b^2} = \sqrt{\frac{2c}{1+c}}\ . \tag{5.8}$$

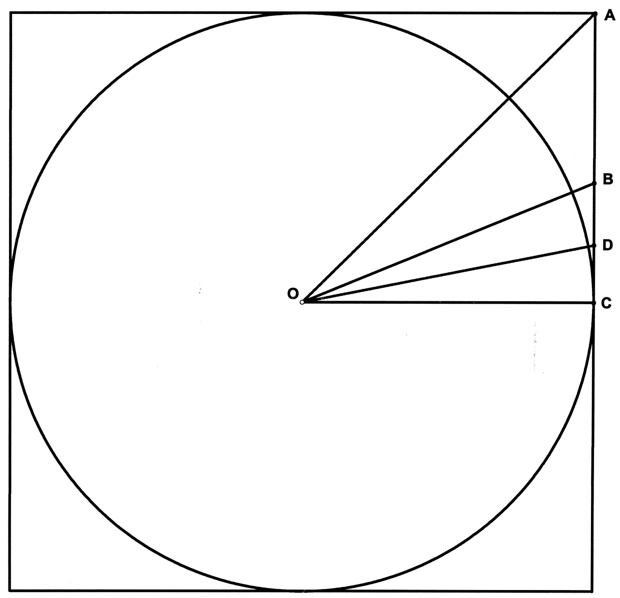


Figure 5: ∠AOC = 45°, ∠BOC = 22.5°, ∠DOC = 11.25°, |AC| = b_1 = |OC| = 1

The relation between a and each of b, c, d is more complex and requires solving a polynomial of third degree. Respectively, from (5.5) we obtain,

$$b^3 + ab^2 + b - a = 0, \tag{5.9}$$
$$c^3 - c^2 - a^2c - a^2 = 0, \tag{5.10}$$
$$d^6 - (a^2+1)d^4 + 4a^2d^2 - 4a^2 = 0. \tag{5.11}$$

We solve (5.9-5.11) and obtain the following real zeros,

$$\begin{aligned}
c &= \frac{1}{3}(18a^2 + 1 + 3a\sqrt{3}\sqrt{1 + 11a^2 - a^4})^{(1/3)} + \frac{1/3 + a^2}{(18a^2 + 1 + 3a\sqrt{3}\sqrt{1 + 11a^2 - a^4})^{(1/3)}} + \frac{1}{3}\ , \\
b &= \frac{1}{3}(18a - a^3 + 3\sqrt{3}\sqrt{1 + 11a^2 - a^4})^{(1/3)} + \frac{a^2/3 - 1}{(18a - a^3 + 3\sqrt{3}\sqrt{1 + 11a^2 - a^4})^{(1/3)}} - \frac{1}{3}a\ , \\
d^2 &= \frac{1}{3}(1 + a^4 + a^2(a^2 - 13)(a^2 - 3) + 6a\sqrt{3}\sqrt{1 + 11a^2 - a^4})^{(1/3)} \\
&+ \frac{(a^2 - 5)^2/3 - 8}{(1 + a^4 + a^2(a^2 - 13)(a^2 - 3) + 6a\sqrt{3}\sqrt{1 + 11a^2 - a^4})^{(1/3)}} + \frac{1}{3}(a^2 + 1)\ .
\end{aligned}$$

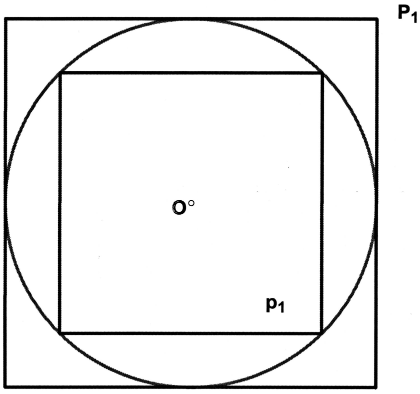


Figure 6: Circumscribed and inscribed squares, P_1, p_1

To the authors' knowledge, the above results are new. In terms of trigonometric functions we have the following identities, (fig. 4),

$$\begin{aligned} a &= \tan\left(\frac{\pi}{2^{n+1}}\right) - \tan\left(\frac{\pi}{2^{n+2}}\right) \\ b &= \tan\left(\frac{\pi}{2^{n+2}}\right) , \\ c &= \sec\left(\frac{\pi}{2^{n+1}}\right) , \\ d &= \sec\left(\frac{\pi}{2^{n+2}}\right) . \end{aligned}$$

In the next section we derive the recursion (2.7) using geometry. We relate the functional form of (2.7) to a Lipschitz mapping and show 'closeness' of b_n to b_{n-1} and s_n to s_{n-1}. We also analyze the convergence of the iterative algorithm in terms of the number of significant digits and show convergence is linear.

6. Numerical Algorithm and Convergence properties

By inspection of fig. 4, we can derive the recursion (2.7) based on (5.1) and the Pythagorean formula. Define

$$b_{n-1} = b_n + a_n \ , \tag{6.1}$$

$$c_n = \frac{a_n}{b_n} = \sqrt{1 + b_{n-1}^2} \ , n = 2, 3, 4, \dots \ . \tag{6.2}$$

Eliminating a_n from (6.1, 6.2) yields

$$b_n = \frac{b_{n-1}}{1 + \sqrt{1 + b_{n-1}^2}} . \tag{6.3}$$

Replacing n by $n+1$ we obtain (2.7). By definition,

$$b_n = \tan\frac{\pi}{2^{n+1}} \ , n = 1, 2, \dots \ . \tag{6.4}$$

Employing (6.4) yields

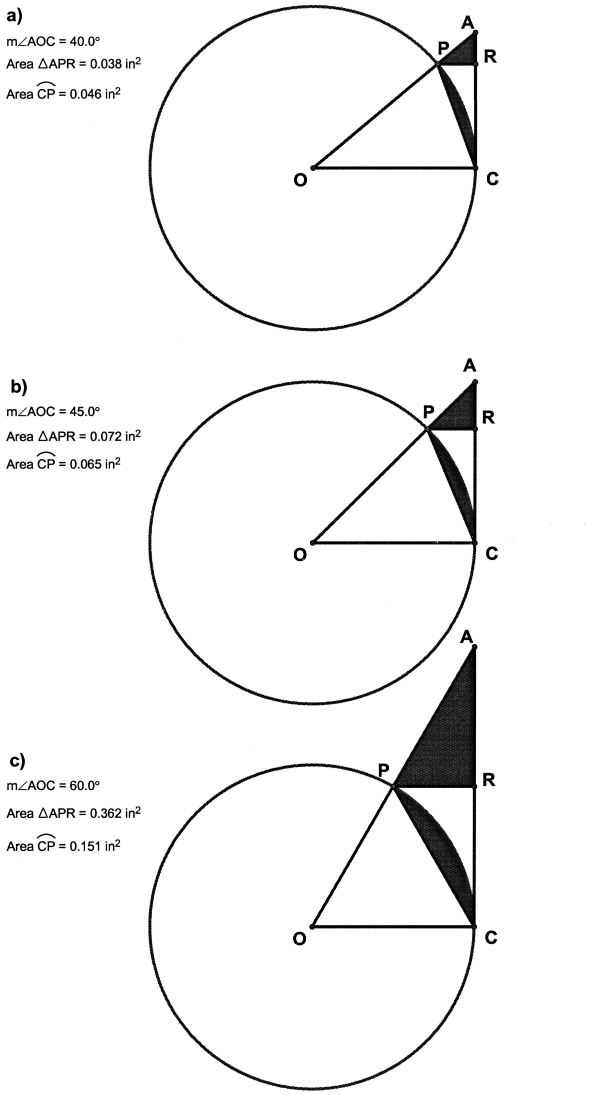


Figure 7: Areas $\triangle APR$ and $\overset{\frown}{CP}$ for a) θ = 40°; b) θ = 45°; and c) θ = 60°.

Lemma 6.1.

$$|b_n - b_{n-1}| = |\tan\frac{\pi}{2^{n+1}} - \tan\frac{\pi}{2^n}| \le \frac{\sqrt{2}(\sqrt{2}-1)}{2^{n-2}}\ , n \ge 2. \tag{6.5}$$

Proof. We see that

$$f(x) = \frac{x}{1+\sqrt{1+x^2}}$$

is Lipschitz mapping with constant $1/2$ since

$$f'(x) = \frac{1}{(1+\sqrt{1+x^2})\sqrt{1+x^2}} \le \frac{1}{2}\ as\ \ x \to 0^+\ .$$

Thus,

$$|b_n - b_{n-1}| \leq \frac{1}{2}|b_{n-1} - b_{n-2}| \leq \ldots \leq \frac{1}{2^{n-2}}|b_2 - b_1| \leq \frac{\sqrt{2}(\sqrt{2}-1)}{2^{n-2}} \approx \frac{2.343146}{2^n} .$$

□

Employing (6.4) yields

Lemma 6.2.

$$\frac{|s_n - s_{n-1}|}{2} = |\sin\frac{\pi}{2^{n+1}} - \sin\frac{\pi}{2^n}| \leq \frac{\sqrt{2}(\sqrt{2}-1)}{2^{n-2}} \ , n \geq 2. \tag{6.6}$$

Proof. We have that,

$$|s_n - s_{n-1}|/2 \leq |b_n - b_{n-1}| .$$

since $\sin x \leq \tan x$ and $\cos x \leq sec^2 x$ near $x = 0$. □

From the Taylor series for tan and sin, cos we obtain, respectively,

$$\begin{aligned} area(P_n) &= 2^{n+1}\tan\left(\frac{\pi}{2^{n+1}}\right) \\ &= \pi\left(1 + \left(\frac{\pi}{2^{n+1}}\right)^2\frac{1}{3} + \left(\frac{\pi}{2^{n+1}}\right)^4\frac{2}{15} + \left(\frac{\pi}{2^{n+1}}\right)^6\frac{17}{315} + \ldots\right) . \end{aligned} \tag{6.7}$$

$$\begin{aligned} area(p_n) &= 2^{n+1}\sin\left(\frac{\pi}{2^{n+1}}\right)\cos\left(\frac{\pi}{2^{n+1}}\right) = 2^n\sin\left(\frac{\pi}{2^n}\right) \\ &= \pi\left(1 - \left(\frac{\pi}{2^n}\right)^2\frac{1}{6} + \left(\frac{\pi}{2^n}\right)^4\frac{1}{120} - \left(\frac{\pi}{2^n}\right)^6\frac{1}{5040} + \ldots\right) . \end{aligned} \tag{6.8}$$

From (6.6-6.7) for sufficiently large $n_0, \forall n > n_0$,

$$\left(\frac{\pi}{2^{n+2}}\right)^2 < \frac{area(P_n) - \pi}{\pi} < \left(\frac{\pi}{\sqrt{2}\cdot 2^{n+1}}\right)^2 . \tag{6.9}$$

$$\left(\frac{\pi}{\sqrt{2}\cdot 2^{n+1}}\right)^2 < \frac{|area(p_n) - \pi|}{\pi} < \left(\frac{\pi}{2^{n+1}}\right)^2 . \tag{6.10}$$

The relative error in approximating π by $area(P_n)$ is from (6.7) of magnitude $K10^{-0.60206\cdot n}$ where $2^n = 10^{0.30103\cdot n}$, $\pi^2/16 < K < \pi^2/8$. The number of significant digits in approximating π by $area(P_n)$ is $O(0.6 \cdot n)$, similarly for $area(p_n)$. The accuracy can be improved to $O(1.2 \cdot n)$ by taking a linear combination of (6.6, 6.7), that is, we have,

$$\left|\frac{area(p_n) + 2\cdot area(P_n)}{3} - \pi\right| \approx 10^{-1.2n} . \tag{6.11}$$

We outline the numerical algorithm that approximates π :
.1 choose $b_1 = 1.0, s_1 = \sqrt{2}$ and a (large enough) positive integer n_0 .

.2 iterate

$$b_n = \frac{b_{n-1}}{1+\sqrt{1+b_{n-1}^2}} \;, n = 2,3,4,\dots\ n_0 \;,$$
$$s_n = \sqrt{2-\sqrt{4-s_{n-1}^2}} \;, n = 2,3,4,\dots\ n_0 - 1 \;.$$

.3 compute

$$\frac{2^{n_0+1}\cdot 2\cdot b_{n_0} + 2^{n_0-1} s_{n_0-1}}{3} \;.$$

Example 1. Letting $n_0 = 599$ and using 600 (MAPLE) significant digits in (6.11) yields,

$$|\frac{2^{600}\cdot 2\cdot b_{599} + 2^{598} s_{598}}{3} - \pi| < 10^{-720}.$$

Example 2. Letting $n = 1998$ with 2000 significant digits in lemmas 6.1, 6.2 obtains,

$$|b_{1999} - b_{1998}| - |s_{1999} - s_{1998}|/2 \approx 0.573610^{-1803} \;.$$

Example 3. Fig. 7 (Geometer's Sketchpad), with several area calculations, is directly related to lemma 4.2.
The structure of $b_n,\ n = 1,2,..,5$ is shown in,
Example 4. ($b_1 = 1,\ b_2 = \frac{1}{1+\sqrt{2}} = -1+\sqrt{2}$.)

$$b_3 = \frac{1}{(1+\sqrt{2})\left(1+\sqrt{1+\frac{1}{(1+\sqrt{2})^2}}\right)} = -1-\sqrt{2}+\sqrt{2}\sqrt{2+\sqrt{2}}$$

$$b_4 = \frac{1}{(1+\sqrt{2})\left(1+\sqrt{1+\frac{1}{(1+\sqrt{2})^2}}\right)\left(1+\sqrt{1+\frac{1}{(1+\sqrt{2})^2\left(1+\sqrt{1+\frac{1}{(1+\sqrt{2})^2}}\right)^2}}\right)}$$

$$= \quad -1-\sqrt{2}-\sqrt{2}\sqrt{2+\sqrt{2}}+\sqrt{2}\sqrt{2+\sqrt{2}}\sqrt{2+\sqrt{2+\sqrt{2}}}\,,$$

$$b_5 \quad = \quad \frac{1}{(1+\sqrt{2})\left(1+\sqrt{1+\frac{1}{(1+\sqrt{2})^2}}\right)\left(1+\sqrt{1+\frac{1}{(1+\sqrt{2})^2\left(1+\sqrt{1+\frac{1}{(1+\sqrt{2})^2}}\right)^2}}\right)}\times$$

$$\frac{1}{1+\sqrt{1+\frac{1}{(1+\sqrt{2})^2\left(1+\sqrt{\frac{1}{(1+\sqrt{2})^2}}\right)^2\left(1+\sqrt{1+\frac{1}{(1+\sqrt{2})^2\left(1+\sqrt{1+\frac{1}{(1+\sqrt{2})^2}}\right)^2}}\right)^2}}}$$

$$= \quad -1-\sqrt{2}-\sqrt{2}\sqrt{2+\sqrt{2}}-\sqrt{2}\sqrt{2+\sqrt{2}}\sqrt{2+\sqrt{2+\sqrt{2}}}$$

$$+ \quad \sqrt{2}\sqrt{2+\sqrt{2}}\sqrt{2+\sqrt{2+\sqrt{2}}}\sqrt{2+\sqrt{2+\sqrt{2+\sqrt{2}}}}\,.$$

We have shown the practical use of the Taylor series in a classical approximation problem.

References

[1] http://mathworld.wolfram.com/ArchimedesRecurrenceFormula.html

[2] Alfred Baker. *Geometry for Schools Theoretical*, W. J. Gage and Company, Limited, Toronto, 1904.

[3] P. Beckmann. *A History of Pi*, Barnes & Noble Books, New York, 1993.

[4] William B. Berlinghoff and Fernando Q. Gouvêa. *Math Through the Ages A Gentle History for Teachers and Others*, Oxford House Publishers, 2002.

[5] B. C. Berndt and Robert A. Rankin, Eds.*RAMANUJAN: Essays and Surveys*, American Mathematical Society, London Mathematical Society, 2001.

[6] http://mathworld.wolfram.com/Brent-SalaminFormula.html

[7] http://mathworld.wolfram.com/BBPFormula.html

[8] E. J. Borowski and J. M. Borwein. *Dictionary of Mathematics*, Collins Reference, Glasgow, 1989.

[9] Jonathan M. Borwein and Peter B. Borwein. *PI and the AGM, A Study in Analytic Number Theory and Computational Complexity*, Canadian Mathematical Society Series of Monographs and Advanced Texts, John Wiley and Sons, 1987.

[10] David M. Burton. *The History of Mathematics An Introduction, 4th Ed.*, WCB, McGraw-Hill, 1999.

[11] Pierre Eymard and Jean-Pierre Lafon. *The Number* π, Amercian Mathematical Society, New York, 2004.

[12] George Grossman *Symmetry and Approximating* π, pre-print from presentation, Mathematics and Symmetry, 2004 Mathematics and Statistics Conference, Miami University, Ohio.

[13] G.H. Hardy, P.V. Seshu Aiyar and B.M. Wilson, Eds., *Modular Equations and Approximations to Pi*, pp. 23-39, in Collected Papers of Srinivasa Ramanaujan, AMS Chelsea Publishing, American Mathematical Society, 2000.

[14] T. L. Heath, Ed. *The Works of Archimedes*, Dover Publications, New York, 2002.

[15] Felix Klein. *Famous problems of Elementary Geometry*, Dover Publications Inc., New York, 1956.

[16] M. J. Maron. *Numerical Analysis A Practical Approach Snd. Ed.*, MacMillan Publisheing Co., New York, 1987.

[17] Joseph Rotman. *Journey into Mathematics, An Introduction to Proofs*, Prentice Hall, New Jersey, 1998.

[18] V.S. Varadajaran. *Algebra in Ancient and Modern Times*, Mathematical World, Vol 12, American Mathematical Society, Hindustan Book Agency, 1998.

[19] http://mathworld.wolfram.com/Arithmetic-GeometricMean.html

JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS,VOL.5,NO.3,197-202,2007,COPYRIGHT 2007 EUDOXUS PRESS,LLC

ON THE RATE OF CONVERGENCE OF MODIFIED SZASZ-MIRAKYAN OPERATORS

Vijay Gupta
School of Applied Sciences, Netaji Subhas Institute of Technology
Sector 3 Dwarka New Delhi-110075, India
vijaygupta2001@hotmail.com

Ram N. Mohapatra
Department of Mathematics, University of Central Florida
Orlando, FL 32816, USA
ramm@mail.ucf.edu

Abstract: In the present paper, we introduce the Bezier variant of modified Szasz-Mirakyan operators for functions of bounded variation. Our result extend and improve the results due to Sahai and Prasad [Publ. Inst. Math. N.S.53(1993),73-80] and Gupta and Pant [J. Math. Anal Appl. 233(1999),476-483].

Keywords: Rate of convergence, Szasz-Mirakyan operators, Bounded variation.
Subject Classification: 41A25, 41A30.

1. INTRODUCTION

Modified Szasz-Mirakyan operators [4] on $L_1[0,\infty)$ are defined as

$$M_n(f;x) = n\sum_{k=0}^{\infty} p_{n,k}(x) \int_0^{\infty} f(t)p_{n,k}(t)\ \ dt,\ \ 0 \le x < \infty, \tag{1}$$

where $p_{n,k}(x) = \exp(-nx)\frac{(nx)^k}{k!}$.

The rate of convergence for these operators for functions of bounded variation was studied recently by Gupta and Pant [2], they also improved and corrected the results of [1] and [5].We now introduce the Bezier variant of these modified Szasz operators as

$$M_{n,\alpha}(f;x) = n\sum_{k=0}^{\infty} Q_{n,k}^{(\alpha)}(x) \int_0^{\infty} f(t)p_{n,k}(t)\ \ dt,\ \ 0 \le x < \infty, \tag{2}$$

where $Q_{n,k}^{(\alpha)}(x) = J_{n,k}^{\alpha}(x) - J_{n,k+1}^{\alpha}(x)$ and $\sum_{j=k}^{\infty} p_{n,j}(x) = J_{n,k}(x)$ is the Szasz-Mirakyan basis function. Obviously $M_{n,\alpha}(f;x) = 1$ and particularly when $\alpha = 1$, the operators (2) reduce to the operators (1) studied in [1]- [5] etc. For further properties of $Q_{n,k}^{(\alpha)}(x)$ and $J_{n,k}(x)$, we refer the readers to [6].

In the present paper, we obtain the rate of convergence for these generalized Szasz-Mirakyan-Bezier operators $M_{n,\alpha}(f;x)$ on functions of bounded variation.

2. AUXILIARY RESULTS

In this section we give certain results, which are necessary to prove the main result.

It is well known that the basis function $p_{n,k}$ corresponds with the Poisson distribution in the probability theory. Using Berry Esseen theorem Gupta and Pant [2] recently obtained the inequality

$$p_{n,k} \leq \frac{32x^2+24x+5}{2\sqrt{nx}}, x \in (0,\infty).$$

In [4] Gupta et al. improved this estimate and obtained the inequality

$$p_{n,k} \leq \frac{\sqrt{3}}{2\sqrt{\pi nx}}, x \in (0,\infty).$$

where $\lim_{k\to\infty} \exp(-nx)\frac{(nx)^k}{k!} \leq \frac{1}{\sqrt{2\pi nx}}$.
Zeng [6,Lemma 3] estimated a more sharp estimate, using his result we have the following lemma:
Lemma 1. For all $x > 0$ and $n, k \in N$ there holds

$$Q_{n,k}^{(\alpha)}(x) \leq \alpha p_{n,k}(x) < \frac{\alpha}{\sqrt{\pi nx}}.$$

Lemma 2. Let the m-th order moment be defined by

$$T_{n,m}(x) \equiv M_{n,1}((t-x)^m, x) = n\sum_{k=0}^{\infty} p_{n,k}(x)\int_0^{\infty} p_{n,k}(t)(t-x)^m dt$$

then we have $T_{n,0}(x) = 1, T_{n,1}(x) = \frac{1}{n}$ and

$$nT_{n,m+1}(x) = xT_{n,m}^{(1)}(x) + (m+1)T_{n,m}(x) + 2mxT_{n,m-1}(x), m \geq 1.$$

From the above recurrence relation, we have

$$T_{n,2}(x) = \frac{2}{n}\left(x+\frac{1}{n}\right).,$$

For $n \geq 2$, we have

$$T_{n,2}(x) \equiv M_{n,1}((t-x)^2, x) \leq \frac{1+2x}{n}.$$

Next let us assume that $K_{n,\alpha}(x,t) = n\sum_{k=0}^{\infty} Q_{n,k}^{(\alpha)}(x)p_{n,k}(t)$, and $\beta_{n,\alpha}(x,y) = \int_0^y K_{n\alpha}(x,t)dt$. Also it is obviously seen that

$$\beta_{n,\alpha}(x,\infty) = \int_0^{\infty} K_{n\alpha}(x,t)dt = 1.$$

Lemma 3. Let $n \geq 2$, then

$$\int_0^y K_{n,\alpha}(x,t)dt \leq \frac{\alpha(1+2x)}{n}\frac{1}{(x-y)^2} \quad , \quad 0 \leq y < x \tag{3}$$

and

$$\int_z^\infty K_{n,\alpha}(x,t)dt \le \frac{\alpha(1+2x)}{n}\frac{1}{(z-x)^2} \quad , \quad x<z<\infty. \tag{4}$$

Proof. Clearly

$$\int_0^y K_{n,\alpha}(x,t)dt \le \int_0^y K_{n,\alpha}(x,t)\frac{(x-t)^2}{(x-y)^2}dt \le \frac{1}{(x-y)^2}M_{n,\alpha}((t-x)^2;x)$$

$$\le \frac{\alpha M_{n,1}((t-x)^2,x)}{(x-y)^2} \le \frac{\alpha(1+2x)}{n(x-y)^2}.$$

The proof of (4) is similar.

3. MAIN RESULT

In this section we prove the following main theorem:

Theorem 1. Let f be a function of bounded variation on every finite subinterval of $[0,\infty)$ and $V_a^b(g_x)$ be the total variation of g_x on $[a,b]$. Also let $\alpha \ge 1$ and $f(t) = O(e^{\gamma t}), t\to\infty$ for some $\eta>0$ then for every $x\in(0,\infty)$ and $n\ge \max\{2,4\eta\}$, we have

$$\left|M_{n,\alpha}(f,x) - \frac{1}{\alpha+1}[f(x+)+\alpha f(x-)]\right| \le \frac{\alpha}{\sqrt{\pi n x}}|f(x+)-f(x-)|$$

$$+\left(\frac{3\alpha(1+2x)}{nx^2}+\frac{1}{n}\right)\sum_{k=1}^n V_{x-x/\sqrt{k}}^{x+x/\sqrt{k}}(g_x) + \frac{\alpha e^{2\eta x}}{x}\sqrt{\frac{2(1+2x)}{n}} + \frac{\alpha e^{\eta x}(1+2x)}{nx^2}, \tag{5}$$

where

$$g_x(t) = \begin{cases} f(t)-f(x^-), \; 0\le t<x; \\ 0, \quad t=x; \\ f(t)-f(x^+), \; x<t<\infty. \end{cases}$$

Proof. Following [7], we have

$$\left|M_{n,\alpha}(f,x) - \frac{1}{\alpha+1}[f(x+)+\alpha f(x-)]\right|$$

$$\le |M_{n,\alpha}(g_x,x)| + \frac{1}{2}\left|M_{n,\alpha}(sign(t-x),x)+\frac{\alpha-1}{\alpha+1}\right|.|f(x+)-f(x-)| \tag{6}$$

First

$$M_{n,\alpha}(sign(t-x),x) = \int_x^\infty K_{n,\alpha}(x,t)dt - \int_0^x K_{n,\alpha}(x,t)dt$$

$$= -1+2\int_x^\infty K_{n,\alpha}(x,t)dt$$

Now using the fact that $n\int_x^\infty p_{n,k}(t)dt = \sum_{j=0}^k p_{n,j}(x)$ for $x\in(0,\infty)$, we have

$$M_{n,\alpha}(sign(t-x),x) = -1+2n\sum_{k=0}^\infty Q_{n,k}^{(\alpha)}(x)\int_x^\infty p_{n,k}(t)dt$$

$$= -1 + 2\sum_{k=0}^{\infty} Q_{n,k}^{(\alpha)}(x) \sum_{j=0}^{k} p_{n,j}(x)$$

$$= -1 + 2\sum_{j=0}^{\infty} p_{n,j}(x) \sum_{k=j}^{\infty} Q_{n,k}^{(\alpha)}(x) = -1 + 2\sum_{j=0}^{\infty} p_{n,j}(x) J_{n,j}^{\alpha}(x).$$

Thus we have

$$M_{n,\alpha}(sign(t-x),x) + \frac{\alpha-1}{\alpha+1} = 2\sum_{j=0}^{\infty} p_{n,j}(x) J_{n,j}^{\alpha}(x) - \frac{2}{\alpha+1}\sum_{j=0}^{\infty} Q_{n,j}^{(\alpha+1)}(x)$$

Since $\sum_{j=0}^{\infty} Q_{n,j}^{(\alpha+1)}(x) = 1$, by mean value theorem, it follows

$$Q_{n,j}^{(\alpha+1)}(x) = J_{n,j}^{\alpha+1}(x) - J_{n,j+1}^{\alpha+1}(x) = (\alpha+1)p_{n,j}(x)\gamma_{n,j}^{\alpha}(x),$$

where $J_{n,j+1}^{\alpha}(x) < \gamma_{n,j}^{\alpha}(x) < J_{n,j}^{\alpha}(x)$. Therefore

$$\left| M_{n,\alpha}(sign(t-x),x) + \frac{\alpha-1}{\alpha+1} \right| = 2\sum_{j=0}^{\infty} p_{n,j}(x)(J_{n,j}^{\alpha}(x) - \gamma_{n,j}^{\alpha}(x))$$

$$\leq 2\sum_{j=0}^{\infty} p_{n,j}(x)(J_{n,j}^{\alpha}(x) - J_{n,j+1}^{\alpha}(x)) \leq 2\alpha \sum_{j=0}^{\infty} p_{n,j}^2(x),$$

where we have used the inequality $b^{\alpha} - a^{\alpha}, \alpha((b-a), 0 \leq a, b \leq 1$ and $\alpha \geq 1$. Applying Lemma 1, we get

$$\left| M_{n,\alpha}(sign(t-x),x) + \frac{\alpha-1}{\alpha+1} \right| \leq \frac{2\alpha}{\sqrt{\pi n x}}, x \in (0,\infty). \tag{7}$$

Next we estimate $M_{n,\alpha}(g_x, x)$ as follows:

$$M_{n,\alpha}(g_x,x) = \int_0^{\infty} K_{n,\alpha}(x,t) g_x(t) dt = \left(\int_{I_1} + \int_{I_2} + \int_{I_3} \right) K_{n,\alpha}(x,t) g_x(t) dt$$

$$=: E_1 + E_2 + E_3, \tag{8}$$

where $I_1 = [0, x - x/\sqrt{n}]$, $I_2 = [x - x/\sqrt{n}, x + x/\sqrt{n}]$ and $I_3 = [x + x/\sqrt{n}, \infty)$. We start with the estimate of E_2. For $t \in [x - x/\sqrt{n}, x + x/\sqrt{n}]$, we have

$$|E_2| \leq V_{x-x/\sqrt{n}}^{x+x/\sqrt{n}}(g_x) \leq \frac{1}{n}\sum_{k=1}^{n} V_{x-x/\sqrt{k}}^{x+x/\sqrt{k}}(g_x). \tag{9}$$

We now estimate E_1, writing $y = x - x/\sqrt{n}$ and using Lebesgue-Stieltjes integration by parts, we have

$$E_1 = \int_0^y g_x(t) d_t(\beta_{n,\alpha}(x,y)) = g_x(y+)\beta_{n,\alpha}(x,y) - \int_0^y \beta_{n,\alpha}(x,y) d_t(g_x(t)).$$

Using Eq. (3) of Lemma 3, we have

$$|E_1| \leq V_{y+}^{x} \frac{\alpha(1+2x)}{n(t-y)^2} + \frac{\alpha(1+2x)}{n} \int_0^y \frac{1}{(x-t)^2} d_t(-V_t^x(g_x)).$$

Integrating by parts the last term, we have after simple computation

$$|E_1| \leq \frac{\alpha(1+2x)}{n} \left[\frac{V_0^x(g_x)}{x^2} + 2\int_0^y \frac{V_t^x(g_x)}{(x-t)^3} \right].$$

Now replacing the variable y in the last integral by $x - x/\sqrt{n}$, we get E_2. For $t \in [x - x/\sqrt{n}, x + x/\sqrt{n}]$, we have

$$|E_1| \leq \frac{2\alpha(1+2x)}{nx^2} \sum_{k=1}^{n} V_{x-x/\sqrt{k}}^{x}(g_x). \tag{10}$$

Finally we estimate E_3. Writing $z = x + x/\sqrt{n}$, we have

$$E_3 = \int_z^\infty K_{n,\alpha}(x,t) g_x(t) dt = \int_z^\infty g_x(t) d_t(\beta_{n,\alpha}(x,t)).$$

We define $R_{n,\alpha}(x,t)$ on $[0, 2x]$ as

$$R_{n,\alpha}(x,t) = \begin{cases} 1 - \beta_{n,\alpha}(x,t), & 0 \leq t < 2x; \\ 0, \quad t = 2x. \end{cases}$$

Therefore

$$E_3 = \int_z^{2x} g_x(t) d_t(R_{n,\alpha}(x,t)) dt - g_x(2x) \int_{2x}^\infty K_{n,\alpha}(x,u) du + \int_z^\infty g_x(t) d_t(\beta_{n,\alpha}(x,t))$$

$$=: E_{31} + E_{32} + E_{33}. \tag{11}$$

By partial integration we have

$$E_{31} = g_x(z-) R_{n,\alpha}(x, z-) + \int_z^{2x} d_t(g_x(t)) \hat{R}_{n,\alpha}(x,t)$$

where $\hat{R}_{n,\alpha}(x,t)$ is the normalized form of $R_{n,\alpha}(x,t)$. Since and $R_{n,\alpha}(x, z-) = R_{n,\alpha}(x,z)$, we have

$$E_{31} = g_x(z-) V_x^{z-}(g_x) R_{n,\alpha}(x,z) + \int_z^{2x} d_t(V_x^t(g_x)) \hat{R}_{n,\alpha}(x,t)$$

Now using Lemma 3 and the fact that $\hat{R}_{n,\alpha}(x,t) \leq R_{n,\alpha}(x,t)$ on $[0, 2x]$, we get

$$|E_{31}| \leq V_x^{z-}(g_x) \frac{\alpha(1+2x)}{n(z-x)^2} + \frac{\alpha(1+2x)}{n} \int_z^{2x-} \frac{1}{(x-t)^2} d_t(V_x^t(g_x)) + \frac{1}{2} V_x^{2x-}(g_x) \int_{2x}^\infty K_{n,\alpha}(x,u) du$$

$$| \leq V_x^{z-}(g_x) \frac{\alpha(1+2x)}{n(z-x)^2} + \frac{\alpha(1+2x)}{n} \int_z^{2x-} \frac{1}{(x-t)^2} d_t(V_x^t(g_x)) + \frac{1}{2} V_x^{2x-}(g_x) \frac{\alpha(1+2x)}{nx^2}.$$

Thus, by replacing the variable in the above integral by $x + x/\sqrt{n}$, we have

$$|E_{31}| \leq \frac{2\alpha(1+2x)}{nx^2} \sum_{k=1}^{n} V_x^{x+x/\sqrt{k}}(g_x). \tag{12}$$

Again using (4) and arguing similarly, we have

$$|E_{32}| \leq g_x(2x) \frac{\alpha(1+2x)}{nx^2} \leq \frac{\alpha(1+2x)}{nx^2} \sum_{k=1}^{n} V_x^{x+x/\sqrt{k}}(g_x). \tag{13}$$

Finally for $n > 4\eta$ and proceeding along the lines of [2], we get

$$|E_{33}| \leq \frac{\alpha e^{2x\eta} \sqrt{(1+2x)}}{x\sqrt{n}} + \frac{\alpha(1+2x) e^{x\eta}}{nx^2}. \tag{14}$$

Collecting the estimates of (6) to (14), we get the required result(5).

This completes the proof of the theorem.

References

[1] V. Gupta and P. N. Agrawal, An estimate on the rate of convergence for modified Szasz-Mirakyan operators of functions of bounded variation, Publ. Inst. Math. Beograd (N.S.) 49(63)(1991), 97-103.

[2] V. Gupta and R. P. Pant, Rate of convergence for the modified Szasz-Mirakyan operators on function of bounded variation, J. Math. Anal. Appl. 233 (1999), 476-483.

[3] V. Gupta, P. Gupta and M. Rogalski, Improved rate of convergence for the modified Szasz-Mirakyan operators, Approx. Theory and its Appl. 16:3 (2000), 94-99

[4] S. M. Mazhar and V. Totik, Approximation by modified Szasz operators, Acta Sci. Math. (Szeged) 49(1985), 257-269.

[5] A. Sahai and G. Prasad, On the rate of convergence for modified Szasz-Mirakyan operators on functions of bounded variation, Publ. Inst. Math. Beograd (N.S.) 53(67)(1993), 73-80.

[6] X. M. Zeng, On the rate of convergence of the generalized Szasz type operators for functions of bounded variation, J. Math. Anal. Appl. 226 (1998), 309-325.

[7] X. M. Zeng and W. Chen, On the rate of convergence of the generalized Durrmeyer type operators for functions of bounded variation, J. Approx. Theory 102(2000) 1-12.

JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS,VOL.5,NO.3,203-211,2007,COPYRIGHT 2007 EUDOXUS PRESS,LLC

An Unsupervised Sequential Algorithm for Pattern Recognition

Iuliana Florentina Iatan,

Department of Mathematics and Informatics,
Technical University of Civil Engineering, Bucharest, Romania,
e-mail iuliafi@yahoo.com

Abstract

It may sometimes happen that we do not have all the data points at once, but rather obtain them gradually over a period of time. The sequential K-means procedure offers us the flexibility of updating the means as new data points arrive.

The goal of this paper is to construct the algorithm of the sequential K-means method.

AMS Subject Classification: 62-xx, 62Lxx, 62L10

Keywords: sequential K-means, cluster analysis, criterion function, initial partition

1 Introduction

Cluster analysis is a process used to solve classification problems. Its object is to group data points into clusters so that the degree of association is strong between members of the same cluster and weak between members of different clusters. Thus each cluster describes the class to which its members belong.

Our objective is to find such partitions of a set of n samples which minimize the criterion function (distance, etc.) Due to an overwhelming number of partitions of n elements into K subsets, we can't consider all possibilities.

One approach is iterative optimization described by the basic iterative minimum-squared-error clustering which could be modified to a sequential K-means procedure. The idea is to come up with a reasonable initial

partition and to "move" samples from one group to another if such move improves the value of the criterion function.

We propose to minimize the sum-of-squared-error criterion J_e, written as [1]

$$J_e = \sum_{i=1}^{K} J_i, \tag{1}$$

where an effective error per cluster is defined to be

$$J_i = \sum_{X \in \omega_i} ||X - m_i||^2 \tag{2}$$

and the mean of each cluster is

$$m_i = \frac{1}{n_i} \sum_{X \in \omega_i} X. \tag{3}$$

2 Sequential K-Means

We suppose that a sample $\hat{X}$ currently in class ω_i is moved to ω_j. Then, the mean of the class ω_j, m_j changes to

$$m_j^* = \frac{X_1 + \ldots + X_{n_j} + \hat{X}}{n_j + 1} = \frac{m_j n_j + \hat{X}}{n_j + 1} = \frac{m_j n_j + m_j + \hat{X} - m_j}{n_j + 1},$$

$$m_j^* = m_j + \frac{\hat{X} - m_j}{n_j + 1} \tag{4}$$

and J_j increases to

$$J_j^* = \sum_{X \in \omega_j} ||X - m_j^*||^2 + ||\hat{X} - m_j^*||^2;$$

therefore

$$J_j^* = \sum_{X \in \omega_j} ||X - m_j - \frac{\hat{X} - m_j}{n_j + 1}||^2 + ||\frac{n_j}{n_j + 1}(\hat{X} - m_j^*)||^2 \tag{5}$$

We have

$$\hat{X} - m_j^* = \hat{X} - m_j - \frac{\hat{X} - m_j}{n_j + 1} = \frac{n_j \hat{X} + \hat{X} - \hat{X} + m_j}{n_j + 1} - m_j =$$

$$= \frac{n_j}{n_j+1}\hat{X} + \frac{m_j}{n_j+1} - m_j = \frac{n_j}{n_j+1}\hat{X} + \frac{m_j - m_j n_j - m_j}{n_j+1},$$

namely

$$\hat{X} - m_j^* = \frac{n_j}{n_j+1}(\hat{X} - m_j). \tag{6}$$

Substituting (6) into (5) we shall obtain

$$J_j^* = J_j + \frac{n_j}{n_j+1}||\hat{X} - m_j||^2. \tag{7}$$

Under the assumption that $n_i \neq 1$, the mean of the class ω_i, m_i changes to

$$m_i^* = \frac{X_1 + \ldots + X_{n_i} - \hat{X}}{n_i - 1} = \frac{m_i n_i - \hat{X}}{n_i - 1} = \frac{m_i n_i - m_i + m_i - \hat{X}}{n_i - 1},$$

$$m_i^* = m_i - \frac{\hat{X} - m_i}{n_i - 1} \tag{8}$$

and J_i decreases to

$$J_i^* = J_i - \frac{n_i}{n_i - 1}||\hat{X} - m_i||^2. \tag{9}$$

The transfer of $\hat{X}$ from ω_i to ω_j is advantageous if the decrease in J_i is greater than the increase in J_j. This is the case if

$$\frac{n_i}{n_i - 1}||\hat{X} - m_i||^2 > \frac{n_j}{n_j+1}||\hat{X} - m_j||^2, \tag{10}$$

which happens whenever $\hat{X}$ is closer to m_j than m_i.

If reassignment is profitable, the greatest decrease in sum of squared error is obtained by selecting the cluster for which

$$\frac{n_j}{n_j+1}||\hat{X} - m_j||^2$$

is minimum.

The idea of this method is to come up with a reasonable initial partition and to "move" samples from one group to another if such move improves the value of the criterion function.

We shall present the algorithm of sequential K-means.

Algorithm 1 [I. IATAN]

```
0. FUNCTION SequentialKMeans(N,K,d,et,X,Nmax,eps)
1. n←detnr(K,N,et)
2. m←detmed(N,K,d,X,et,n)
3. s←calcje(K,N,d,X,m,et)
4. it←1
5. i←1
6. WHILE it<Nmax
   6.1. WHILE i<N
      6.1.1. IF n[et[i]]<>1
        THEN
          6.1.1.1. ro←calcro(K,N,d,i,X,m,n,et)
          6.1.1.2. et[i]←pozminimum(ro,K)
          6.1.1.3. n1←detnr(K,N,et)
          6.1.1.4. m1←detmed(N,K,d,X,et,n1)
          6.1.1.5. s1←calcje(K,N,d,X,m1,et)
          6.1.1.6. n←n1
          6.1.1.7. m←m1
      6.1.2. i←i+1
      6.1.3. CONTINUE
   6.2. IF abs(s-s1)<=eps
     THEN
       6.2.1. it←Nmax
   6.3. s←s1
   6.4. it←it+1
   6.5. i←1
   6.6. CONTINUE
7. SequentialKMeans←et
8. RETURN
```

Data input:

- N- the number of samples,
- K- the number of classes,
- d- the space dimensionality,
- et- the vector which contains the labels of the samples from the initial partition,
- X- the matrix having the samples like columns,

- *Nmax*- the number of attempts,
- *eps*- the precision of the method.

Data output: the vector of whose components are the labels of the samples in the final partition.

The function *detnr* returns the number of samples from each of the K classes.

Algorithm 2[I. IATAN]
0. FUNCTION detnr(K,N,et)
1. FOR k=1,2,...,K
   1.1. n[k]←0
   1.2. FOR i=1,2,...,N
      1.2.1. IF et[i]==k
         THEN
            1.2.1.1. n[k]←n[k]+1
      1.2.2. CONTINUE
   1.3. CONTINUE
2. detnr←n
3. RETURN

Data input:

- K- the number of classes,
- N- the number of samples,
- *et*- the vector which contains the labels of the samples.

Data output: the number of samples from each of the K classes.

The function *detmed* calculates the means of the K classes.

Algorithm 3[I. IATAN]
0. FUNCTION detmed(N,K,d,X,et,n)
1. FOR k=1,2,...,K
   1.1. FOR j=1,2,...,d
      1.1.1. m[j,k]←0
      1.1.2. FOR i=1,2,...,N
         1.1.2.1. IF et[i]==k

```
      THEN
        1.1.2.1.1. m[j,k]←m[j,k]+X[j,i]
    1.1.2.2. CONTINUE
  1.1.3. m[j,k]←m[j,k]/n[k]
  1.1.4. CONTINUE
 1.2. CONTINUE
2. detmed←m
3. RETURN
```

Data input:

- N- the number of samples,
- K- the number of classes,
- d- the space dimensionality,
- X- the matrix having the samples like columns,
- *et*- the vector which contains the labels of the samples,
- n- the vector of whose components mean the number of samples from each of the K classes.

Data output: a matrix having the means of the K classes like columns.

The function *calcje* calculates the sum-of-squared-error criterion J_e using (1).

Algorithm 4 [I. IATAN]

```
0. FUNCTION calcje(K,N,d,X,m,et)
1. s←0
2. FOR k=1,2,...,K
  2.1. nor[k]←0
  2.2. FOR i=1,2,...,N
    2.2.1. FOR j=1,2,...,d
      2.2.1.1. IF et[i]==k
        THEN
          2.2.1.1.1. nor[k]←nor[k]+(X[j,i]-m[j,k])*(X[j,i]-m[j,k])
      2.2.1.2. CONTINUE
    2.2.2. CONTINUE
  2.3. s←s+nor[k]
```

2.4. CONTINUE
3. calcje←s
4. RETURN

Data input:

- K- the number of classes,
- N- the number of samples,
- d- the space dimensionality,
- X- the matrix having the samples like columns,
- m- the matrix having the means of the K classes like columns,
- et- the vector which contains the labels of the samples from the respective partition.

Data output: the sum-of-squared-error criterion J_e from (1).

The function *calcro* is used in order to determine:

$$\rho_i = \begin{cases} \frac{n_i}{n_i-1}||X - m_i||^2 \ if \ X \in class \ \omega_i, \\ \\ \frac{n_i}{n_i+1}||X - m_i||^2 \ otherwise, \ for \ i = \overline{1, K}. \end{cases} \tag{11}$$

Algorithm 5 [I. IATAN]
0. FUNCTION calcro(K,N,d,i,X,m,n,et)
1. nor←0
2. FOR j=1,2,...,d
 2.1. nor←nor+(X[j,i]-m[j,et[i]])*(X[j,i]-m[j,et[i]])
 2.2. CONTINUE
3. ro[et[i]]←n[et[i]]/(n[et[i]]-1)*nor
4. FOR h=1,2,...,M
 4.1. IF h<>et[i]
  THEN
   4.1.1. nor←0
   4.1.2. FOR j=1,2,...,d
    4.1.2.1. nor←nor+(X[j,i]-m[j,h])*(X[j,i]-m[j,h])
    4.1.2.2. CONTINUE
   4.1.3. ro[h]←n[h]/(n[h]+1)*nor
 4.2. CONTINUE

5. calcro←ro
6. RETURN

Data input:

- K- the number of classes,
- N- the number of samples,
- d- the space dimensionality,
- i- index of the current vector,
- X- the matrix having the samples like columns,
- m- the matrix having the means of the K classes like columns,
- et- the vector which contains the labels of the samples from the respective partition.

Data output: the vector of whose components are those from (11).

The function *pozminimum* returns the position of the minimum component from a vector.

Algorithm 6

```
0. FUNCTION pozminimum(v,d)
1. t←v[1]
2. FOR i=1,2,...,d
   2.1. IF t>=v[i]
     THEN
       2.1.1. t←v[i]
       2.1.2. j←i
   2.2. CONTINUE
3. pozminimum←j
4. RETURN
```

Data input:

- v- a vector,
- d- the vector dimension.

Data output: the position of the minimum component from a vector.

3 Conclusions

The goal of data clustering, or unsupervised learning, is to discover "natural" groupings in a set of patterns, points, or objects, without prior knowledge of any class labels.

It may sometimes happen that we do not have all the data points at once, but rather obtain them gradually over a period of time. The sequential K-means procedure offers us the flexibility of updating the means as new data points arrive.

Our aim is to design the algorithm of the sequential K-means method.

References

[1] Duda, D.,O., Hart P., E., Stork, D., G., 2001. *Pattern Classification.* John Wiley, second edition.

[2] Enăchescu, D., 2003. *Elements of Statistical Learning. Applications in Data Mining.* Cooperativa Libraria Editrice Università di Padova.

[3] Enăchescu, D., 2004. *Unsupervised Statistical Learning and Data Mining.* Cooperativa Libraria Editrice Università di Padova.

[4] Figueiredo, M., Jain, A., 2002. "Unsupervised Learning of Finite Mixture Models". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(3): 381-396.

[5] Iatan, I., 2006. "A Statistical Supervised Classification Using a Perceptron Criterion Function", *Acta Universitatis Apulensis. Proceedings of the International Conference on Theory and Application of Mathematics and Informatics ICTAMI 2005, Alba- Iulia*, 11: 103-109.

[6] Jain, A., Dubes, R., 1988. *Algorithms for Clustering Data.* Prentice-Hall, Englewood Cliffs, New Jersey 07632.

[7] Nilsson, N., 1997. *Introduction to Machine Learning.* Artificial Intelligence Laboratory, Department of Computer Science Stanford University Stanford, CA 94305.

[8] Ripley, B.,D., 1996. *Pattern Recognition and Neural Networks.* Press Syndicate of the University of Cambridge.

[9] Webb, A., 2002. *Statistical Pattern Recognition.* John Wiley and Sons, N.York, second edition.

JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS,VOL.5,NO.3,213-218,2007,COPYRIGHT 2007 EUDOXUS PRESS,LLC

On the zeros of a class of analytic functions

by

Chrysi G. Kokologiannaki

University of Patras, Department of Mathematics,
26500 Patras, Greece

e-mail address: chrykok@math.upatras.gr

Abstract. The zeros of a class of analytic functions represented by a continued fraction are studied. The results are applied to the zeros of the mixed Bessel function $J_\nu(z) + \mu J_{\nu+1}(z)$, $\mu \in C$ and improve a previously known result.

Keywords: zeros, continued fractions, tridiagonal operators, Bessel functions.

AMS Subject Classification 2000 : 40A15, 47B36, 33A10.

1. Introduction.

We consider the continued fraction

$$\frac{1}{\lambda - b_1 - \frac{a_1^2}{|\lambda - b_2} - \frac{a_2^2}{|\lambda - b_3} - \ldots} \tag{1.1}$$

with a_n, b_n real sequences and $a_n > 0$. To the continued fraction (1.1) corresponds a tridiagonal operator

$$Te_n = a_n e_{n+1} + a_{n-1} e_{n-1} + b_n e_n, \qquad n = 2, 3, .. \tag{1.2}$$

$$Te_1 = a_1 e_2 + b_1 e_1$$

whose definition domain is the set of finite linear combinations of a fixed orthonormal basis e_n, $n = 1, 2, ...$ of a separable Hilbert space H [7].

Let P_N be the orthogonal projection on the subspace H_N spanned by $\{e_1, ..., e_N\}$ and T_N be the truncated tridiagonal operator $T_N = P_N T P_N$. The operator T_N has a discrete spectrum with N eigenvectors in H_N. A real point λ is called a limit point of eigenvalues of T_N if there exists a sequence of eigenvalues $\{\lambda_N\}$ of T_N such that $\lambda_N \to \lambda$ as $N \to \infty$. We denote the set of all such limit points by $\Lambda(T)$ and the spectrum of the operator T by $\sigma(T)$. The following theorem holds:

Theorem 1.1 [7] Let T be self-adjoint operator. Then the continued fraction (1.1) converges to a finite value for every $\lambda \in C - \Lambda(T)$, where C is the set of

complex numbers. The convergence is uniform on compact subsets of $C - \Lambda(T)$ and the value $K(\lambda)$ to which (1.1) converges is given by $K(\lambda) = ((\lambda - T)^{-1}e_1, e_1)$.

In section 2, we'll prove that under certain conditions on T the zeros λ_k of the function $K(\lambda) - \frac{1}{\alpha}$, $\alpha \neq 0$ are the eigenvalues of the operator $T + \alpha P$ where P is the orthogonal projection on the subspace spanned by the element e_1 and give an upper bound of the absolute value of the zeros λ_k in the case where the operator T is bounded. In section 3, as an application, we obtain a lower bound for the absolute value of the complex zeros of the function $J_\nu(z) + \mu J_{\nu+1}(z)$, $\mu \in C$. This inequality restricts the region of the non-existence of the complex zeros of the above function, if they exist, which has been given in [6] and of the special case where $\mu = \pm i$ which has been proved in [9].

2. Zeros of the function $K(\lambda) - \frac{1}{\alpha}$, $\alpha \neq 0$.

Theorem 2.1 Assume that T has a discrete spectrum with eigenvalues ϵ_n such that $|\epsilon_n| \to 0$ and $\sigma(T) = \Lambda(T)$. Then the point $\lambda \neq 0$ is a zero of the function

$$K(\lambda) - \frac{1}{\alpha} \tag{2.1}$$

if and only if λ is an eigenvalue of the operator $T + \alpha P$, where P is the orthogonal projection $Px = (x, e_1)e_1$. Moreover if $\alpha = \alpha_1 + i\alpha_2$, then the zeros λ_k of the function (2.1) satisfy the inequality:

$$|\lambda_k|^2 \leq (\|T\| + |\alpha_1|)^2 + |\alpha_2|^2. \tag{2.2}$$

Proof Let λ be an eigenvalue of $T + \alpha P$, i.e. for $x \neq 0$,

$$(T + \alpha P)x = \lambda x. \tag{2.3}$$

Then

$$(\lambda - T)x = \alpha(x, e_1)e_1. \tag{2.4}$$

Note that the operators T and $T + \alpha P$ have no common eigenvalues. In fact if λ is a common eigenvalue with corresponding eigenvectors x and y respectively, i.e. the equations (2.3) and $Ty = \lambda y$ hold. Then we easily get $\alpha(x, e_1)(e_1, y) = 0$, which is impossible, because if $(x, e_1) = 0$ then from (2.3) since $x = \sum_{n=1}^{\infty}(x, e_n)e_n$ we get that $(x, e_2) = 0$ and going on it follows that all the coefficients (x, e_n) are equal to zero, that is $x = 0$. Similarly it follows that if $(y, e_1) = 0$ then $y = 0$. Thus λ is not an eigenvalue of T and since the spectrum of T is discrete, λ $(\lambda \neq 0)$ is a regular point of T. Thus from equation (2.4) we obtain

$$x = \alpha(x, e_1)(\lambda - T)^{-1}e_1$$

and

$$(x, e_1) = \alpha(x, e_1)((\lambda - T)^{-1}e_1, e_1).$$

Since $(x, e_1) \neq 0$, λ is a solution of the equation

$$((\lambda - T)^{-1}e_1, e_1) = \frac{1}{\alpha}. \tag{2.5}$$

Since $\sigma(T) = \Lambda(T)$ we have [7] $((\lambda - T)^{-1}e_1, e_1) = K(\lambda)$ and λ is a zero of (2.1). This proves the "if" part of the theorem. To prove the "only if" part we work as follows. Write the equation $K(\lambda) = \frac{1}{\alpha}$ as

$$(e_1, e_1) = \alpha((\lambda - T)^{-1}e_1, e_1), \quad \textit{or} \quad (\alpha(\lambda - T)^{-1}e_1 - e_1, e_1) = 0.$$

The last equation means that the elements $y_0 = \alpha(\lambda - T)^{-1}e_1 - e_1$ and e_1 are orthogonal, i.e. $(y_0, e_1) = 0$. Thus $\alpha(\lambda - T)^{-1}P(e_1 + y_0) = e_1 + y_0$, because $Py_0 = 0$. Setting $x = e_1 + y_0$ we find $x \neq 0$ and $(\lambda - T)x = \alpha Px$ or $(T + \alpha P)x = \lambda x$.

Moreover, since the zeros λ_k of the function (2.1) satisfy (2.3), for $\|x_k\| = 1$, from (2.3) we obtain

$$\lambda_k = ((T + \alpha P)x_k, x_k)$$

or

$$\lambda_k = (Tx_k, x_k) + \alpha(Px_k, x_k). \tag{2.6}$$

Since $(Px_k, x_k) = |(x_k, e_1)|^2$ and setting $\alpha = \alpha_1 + i\alpha_2$, (2.6) takes the form:

$$\lambda_k = (Tx_k, x_k) + \alpha_1|(x_k, e_1)|^2 + i\alpha_2|(x_k, e_1)|^2. \tag{2.7}$$

From (2.7) we obtain:

$$|\lambda_k|^2 = ((Tx_k, x_k) + \alpha_1|(x_k, e_1)|^2)^2 + (\alpha_2)^2|(x_k, e_1)|^4 \tag{2.8}$$

and from (2.8) we get the desired inequality (2.2).

Remark 2.1 The operators T and P are self-adjoint so (Tx_k, x_k) and (Px_k, x_k) are real, and from (2.8) it follows that if $\alpha \in R$ then the zeros λ_k of (2.1) are real and if $\alpha \in C$ then the zeros λ_k of (2.1) are complex.

3. An application to the zeros of the function $J_\nu(z) + \mu J_{\nu+1}(z)$, $\mu \in C$.

Let T_ν, $\nu > -1$ be the tridiagonal operator

$$T_\nu e_n = \frac{1}{2\sqrt{(n+\nu)(n+\nu+1)}} e_{n+1} + \frac{1}{2\sqrt{(n+\nu)(n+\nu-1)}} e_{n-1}, \quad n = 1, 2, ... \tag{3.1}$$

This operator has been studied [1, 2, 3, 4] extensively in connection with the zeros $\pm j_{\nu k}$, $k = 1, 2, ...$ of the Bessel function $J_\nu(z)$. For $\nu > -1$ the operator T_ν is self-adjoint and compact [1], so the continued fraction associated with T_ν is

$$\frac{1}{\lambda -} \frac{\frac{1}{4(\nu+1)(\nu+2)}}{\lambda -} \frac{\frac{1}{4(\nu+2)(\nu+3)}}{\lambda -} ... \tag{3.2}$$

and converges to the function $K_\nu(\lambda) = ((\lambda - T_\nu)^{-1} e_1, e_1)$ [7] for every $\lambda \in C - \sigma(T_\nu)$, where C is the set of complex numbers and $\sigma(T_\nu)$ is the spectrum of T_ν, so, $\sigma(T_\nu) = \{0, \pm\frac{1}{j_{\nu,k}}\}$, $k = 1, 2, ...$

In [5] it has been proved that

$$K_\nu(\lambda) = ((\lambda - T_\nu)^{-1} e_1, e_1) = 2(\nu+1)\frac{J_{\nu+1}(1/\lambda)}{J_\nu(1/\lambda)} \tag{3.3}$$

$\nu > -1$ and $\lambda \neq \{0, \pm\frac{1}{j_{\nu k}}\}$, $k = 1, 2, ...$.

Remark 3.1 (i) From (3.3) it follows that the zeros of the equation $K_\nu(\lambda) = 0$ are $\pm\frac{1}{j_{\nu+1,k}}$, $k = 1, 2, ...$ where $j_{\nu+1,k}$ are the zeros of Bessel function $J_{\nu+1}(z)$ and (ii)from (3.3) and using also the known [10] recurrence relation $J_{\nu+2}(z) + J_\nu(z) = \frac{2(\nu+1)}{z} J_{\nu+1}(z)$ it follows that the zeros of the equation $K_\nu(\lambda) = \frac{1}{\lambda}$ are $\pm\frac{1}{j_{\nu+2,k}}$, $k = 1, 2, ...$ where $j_{\nu+2,k}$ are the zeros of Bessel function $J_{\nu+2}(z)$.

Theorem 3.1 The complex zeros z_k of the function

$$J_\nu(z) + \mu J_{\nu+1}(z), \quad \nu > -1, \quad \mu \in C, \tag{3.4}$$

if they exist, satisfy the inequality:

$$|z_k|^2 \geq \left((\frac{1}{j_{\nu,1}} + \frac{|\mu_1|}{2(\nu+1)})^2 + \frac{|\mu_2|^2}{4(\nu+1)^2} \right)^{-1} \tag{3.5}$$

where $j_{\nu,1}$ is the first positive zero of Bessel function $J_\nu(z)$, $\nu > -1$ and μ_1 ,μ_2 is the real and imaginary part of μ respectively.

Proof The equation $J_\nu(z)+\mu J_{\nu+1}(z)=0$ can be written in the form

$$2(\nu+1)\frac{J_{\nu+1}(z)}{J_\nu(z)}=-\frac{2(\nu+1)}{\mu}. \tag{3.6}$$

Setting $z=\frac{1}{\lambda}$ in (3.6) and using (3.3) we get

$$K_\nu(\lambda)=((\lambda-T_\nu)^{-1}e_1,e_1)=-\frac{2(\nu+1)}{\mu}, \tag{3.7}$$

where T_ν is self-adjoint and compact operator [1]. Using Theorem 2.1 for α given by

$$\alpha=-\frac{\mu}{2(\nu+1)}=-\frac{\mu_1}{2(\nu+1)}-\frac{i\mu_2}{2(\nu+1)},$$

it follows that the zeros λ_k of equation (3.7) satisfy

$$\lambda_k=(T_\nu x_k,x_k)-\frac{\mu_1}{2(\nu+1)}|(x_k,e_1)|^2-i\frac{\mu_2}{2(\nu+1)}|(x_k,e_1)|^2. \tag{3.8}$$

or the inequality

$$|\lambda_k|^2\le(\|T_\nu\|+\frac{|\mu_1|}{2(\nu+1)})^2+\frac{|\mu_2|^2}{4(\nu+1)^2}. \tag{3.9}$$

Since $\|T_\nu\|=\frac{1}{j_{\nu,1}}$ [3] and the zeros z_k of the function (3.4) are the inverse of λ_k, from (3.9) follows the desired inequality (3.5).

Remark 3.2 Equating the imaginary parts of (3.8) we obtain

$$Im\lambda_k=-\frac{\mu_2}{2(\nu+1)}|(x_k,e_1)|^2. \tag{3.10}$$

Since the zeros z_k of (3.4) are $\lambda_k=\frac{1}{z_k}=\frac{\overline{z_k}}{|z_k|^2}$, (3.10) becomes

$$Imz_k=\frac{\mu_2}{2(\nu+1)}|(x_k,e_1)|^2|z_k|^2. \tag{3.11}$$

From (3.11) for $\nu>-1$, if $\mu_2>0$ ($\mu_2<0$) then the function (3.4) can not have complex zeros in the lower (upper) half-plane. This result has been proved with a different method in [6]. We mention that using Theorem 3.1 we obtain that the zeros of the function (3.4) if they exist, are located in the upper (lower) half plane

for $\mu_2 > 0$ ($\mu_2 < 0$) and out of the semicircle $(0, ((\frac{1}{j_{\nu,1}} + \frac{|\mu_1|}{2(\nu+1)})^2 + \frac{|\mu_2|^2}{4(\nu+1)^2})^{-1/2})$. This result restricts the region of non-existence of complex zeros of the function (3.4). Especially if $\mu = \pm i$ it follows that the zeros of the function $J_\nu(z) \pm iJ_{\nu+1}(z)$ if they exist, are located in the upper and lower half planes respectively and out of the semicircle $(0, (\frac{1}{2(\nu+1)} + \frac{1}{j_{\nu,1}})^{-1})$. This result agrees with the numerical computation of the zeros of the function $J_\nu(z) \pm iJ_{\nu+1}(z)$ for $\nu = 0, 1, 2, 3, 5$ which have been evaluated in [8,9].

Acknowledgement: The author is grateful to prof. E.K.Ifantis for his helpful suggestions.

References

[1] E.K.IFANTIS, P.D.SIAFARIKAS and C.B.KOURIS, *Conditions for solution of a linear first order differential equation in the Hardy-Lebesque space and applications*, J. Math. Anal. Appl. 104 (1984) 454-466.

[2] E.K.IFANTIS, P.D.SIAFARIKAS and C.B.KOURIS, *Upper bounds for the first zeros of Bessel functions*, J. Comp. Appl. Math. 7 (1984)

[3] E.K.IFANTIS and P.D.SIAFARIKAS, *An inequality relating the zeros of two ordinary Bessel functions*, Applicable Analysis 19 (1985) 251-263.

[4] E.K.IFANTIS and P.D.SIAFARIKAS, *A differential equation for the zeros of Bessel functions*, Applicable Analysis 20 (1985) 269-281.

[5] E.K.IFANTIS and P.D.SIAFARIKAS, *Inequalities involving Bessel and Modified Bessel functions*, J. Math. Anal. Appl. 147, 1, (1990) 214-227.

[6] C.G.KOKOLOGIANNAKI and P.D.SIAFARIKAS, *Non-existence of complex and purely imaginary zeros of a transcendental equation involving Bessel functions*, Zeit. Anal. Annew., 10 (1991), 4, 563-567.

[7] E.K.IFANTIS and P.N.PANAGOPOULOS, *Limit points of eigenvalues of truncated tridiagonal operators* J. Comp. App. Math. 133 (2001) 413-422.

[8] D.A.MACDONALD, *On the computation of zeros of* $J_n(z) - iJ_{n+1}(z) = 0$, Quarterly of Appl. Math., LV, 4, (1997), 623-633.

[9] S.TADEPALLI and C.E.SYNOLAKIS, *Roots of* $J_\gamma(z) \pm iJ_{\gamma+1}(z) = 0$ *and the evaluation of integrals with cylindrical function kernels*, Quarterly of App.Math., LII, 1, (1994), 103-112

[10] G.N.WATSON, *A treatise on the theory of Bessel functions*, Cambridge Univ. Press, 1966.

JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS,VOL.5,NO.3,219-229,2007,COPYRIGHT 2007 EUDOXUS PRESS,LLC

Neural Network for Stochastic Multi-objective Optimization Problems

M. El-Sayed Wahed
Mathematics Department, Faculty of Science,
Zagazig University, Zagazig Egypt
mewahed@yahoo.com

Abstract

In this paper, we propose recurrent neural network for solving a stochastic multi-objective optimization problems with nonlinear constraints. The proposed neural network has a simpler structure and a lower complexity for implementation than the existing neural networks for solving such problems. It is shown here that the proposed neural network is stable in the sense of Lyapunov and globally convergent to an optimal solution. Compared with the existing convergence results, the present results do not require Lipschitz continuity condition on the stochastic multi-objective objective function. Finally, examples are provided to show the applicability of the proposed neural network.

Keywords: Neural network; Stochastic programming; Multi-objective programming; Nonlinear programming.

1 Introduction

Many engineering problems can be solved by transforming the original problems into nonlinearly constrained optimization problems. For example, the least square problem with nonlinear equality constraints can be viewed a basic framework which are widely used for system modeling and design in a variety of applications such as signal and image processing and pattern recognition [1]. In many applications, real-time solutions are usually imperative. One example of such applications in image processing is the solution to the image fusion problem in real-time wireless image transmission [2]. Compared with traditional numerical methods for constrained optimization, the neural network approach has several advantages in real-time applications. First, the structure of a neural network can be implemented effectively using VLSI and optical technologies. Second, neural networks can solve many optimization problems with time-varying parameters. Third, the numerical ordinary differential equation (ODE) techniques can be applied directly to the continuous-time neural network for solving constrained optimization problems effectively. Therefore, neural network method for optimization have been received considerable attention [3]. Many continuous-time neural networks for constrained optimization problems have been developed [17],[18]. At present there exist several neural networks for solving stochastic multi-objective optimization problems with nonlinear constraints [6],[11]. Kennedy and Chua [3] developed a neural network for solving nonlinear programming problems, which fulfills the Kuhn-Tucker optimality condition. Because the Kennedy-Chua network contains a finite penalty parameter, it generates approximate solutions only and has an implementation problem when the penalty parameter is very large. To avoid using penalty parameters, some significant works have been done in recent years. In addition, several projection neural networks for constrained optimization and related problems were developed [11]. In [10], a two-layer neural network for solving nonlinear convex programming problems was studied. This network is shown to be globally convergent to an exact solution under a Lipschitz continuity condition of the objective function. Their network is shown to be globally convergent to an exact solution under a strictly convex condition of the objective function, In this paper, we propose a one-layer neural network for solving such stochastic multiopjective programming problems. Furthermore, we extend the proposed neural network for solving a class of monotone variational inequality with linear equality constraints. The proposed network is shown to be globally convergent to an exact solution within a finite time. Since the low

complexity of neural networks is greatly significant from the viewpoint of computation and implementation, the proposed neural network is an attractive alternative of the existing neural network for deterministic optimization with nonlinear inequality constraints. Finally, simulation results and applied example further confirm the effectiveness of the proposed neural network. The paper is organized as follows. In Section 2, Multi-objective Chance constrained programming technique with a joint constraint. In section 3, Neural networks architectures for nonlinear mixed constraints. In Section 4, A neural network model is proposed to solve deterministic nonlinear optimization problem with equality and non equality constraints. In section 5, gives the numerical example. In section 6. gives the conclusion of this paper.

2 Multi-objective Chance constrained programming technique with a joint constraint [14].

A multi-objective chance constrained programming problem with a joint probability constraint can be stated as

$$\min Z^{(k)}(x) = \sum_{j=1}^{n} c_j^{(k)} x_j, \quad k = 1, 2, ..., K \tag{1}$$

$$\text{Subject to: } Prob[\sum_{j=1}^{n} a_{ij} x_j \geq b_i] \geq 1 - \alpha, i = 1, ..., m, \tag{2}$$

$$x_j \geq 0, \quad j = 1, ..., n, \quad \alpha \in (0, 1). \tag{3}$$

where b_i's are independent normal random variables with known means and variances. Equation (2) is a joint probabilistic constraint and $0 < \alpha < 1$ is a specified probability. We assume that the decision variables x_j's are deterministic.

2.1 Deterministic model

Let the mean and standard deviation of the normal independent random variable b_i be given by $E(b_i)$ and $\sigma(b_i)$ respectively. Equation (2) can be rewritten as

$$\prod_{i=1}^{m} Prob[\sum_{j=1}^{n} a_{ij} x_j \geq b_i] \geq 1 - \alpha \tag{4}$$

It can be simplified as

$$\prod_{i=1}^{m} Prob\left[\frac{b_i - E(b_i)}{\sigma(b_i)} \leq \frac{\sum_{j=1}^{n} a_{ij} x_j - E(b_i)}{\sigma(b_i)}\right] \geq 1 - \alpha \tag{5}$$

where $\left[\frac{b_i - E(b_i)}{\sigma(b_i)}\right]$, $\forall i = 1, ..., m$ is a standard normal variate with zero mean and unit variance. Therefore,

$$\prod_{i=1}^{m} \phi\left[\frac{\sum_{j=1}^{n} a_{ij} x_j - E(b_i)}{\sigma(b_i)}\right] \geq 1 - \alpha \tag{6}$$

where $\phi(.)$ represents the cumulative distribution function of the standard normal random variable. Let

$$\beta_i = \frac{\sum_{j=1}^{n} a_{ij} x_j - E(b_i)}{\sigma(b_i)}, \quad i = 1, ..., m \tag{7}$$

and

$$\phi(\beta_i) = y_i, \quad i = 1, ..., m \tag{8}$$

Then,

$$\prod_{i=1}^{m} y_i \geq 1 - \alpha \tag{9}$$

It is well known that

$$\phi(\beta_i) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\beta_i} \exp(-z^2/2)dz \tag{10}$$

Assuming $z^2/2 = t$, it can be further simplified as

$$\begin{aligned}
\phi(\beta_i) &= \frac{-1}{2\sqrt{\pi}} \int_{\beta_i^2/2}^{\infty} \exp(-t)t^{(1/2-1)}dt && (11)\\
&= \frac{-1}{2\sqrt{\pi}} \left[\int_0^{\infty} \exp(-t)t^{(1/2-1)}dt - \int_0^{\beta_i^2/2} \exp(-t)t^{(1/2-1)}dt \right] && (12)\\
&= \frac{1}{2} \left[\frac{1}{(\Gamma(1/2))^{\gamma}} (\frac{1}{2}, \frac{\beta_i^2}{2}) - 1 \right] && (13)\\
&= \frac{1}{2} \left[p(\frac{1}{2}, \frac{\beta_i^2}{2}) - 1 \right] && (14)
\end{aligned}$$

where

$$\begin{aligned}
p(a,x) &= \frac{\gamma(a,x)}{\Gamma(a)} && (15)\\
&= \frac{1}{\Gamma(a)} \int_0^x \exp(-t)t^{(a-1)}dt, \quad a > 0 && (16)
\end{aligned}$$

is an incomplete gamma function [9] which is monotonic and satisfies the following conditions:

$$P(a,0) = 0 \quad \text{and} \quad P(a,\infty) = 1$$

where

$$\gamma(a,x) = \exp(-x)x^a \sum_{r=0}^{\infty} \frac{\Gamma(a)}{\Gamma(a+1+r)} x^r, \tag{17}$$

So, Equation(10) becomes

$$\exp(-\beta_i^2/2) \sum_{r=0}^{\infty} \frac{\beta_i^{2r+1}}{2^r \Gamma(\frac{3}{2}+r)} = (2y_i+1)\sqrt{2}, \tag{18}$$

This can be rewritten as

$$\sum_{r=0}^{\infty} \frac{\beta_i^{2r+1}}{\prod_{n=0}^{r}(2n+1)} = \sqrt{\frac{\pi}{2}}(2y_i+1)\exp(\beta_i^2/2), \tag{19}$$

we know that

$$\sum_{r=0}^{\infty} \frac{\beta_i^{2r+1}}{\prod_{n=0}^{r}(2n+1)} \tag{20}$$

is convergent for any value of β_i. The series can be expanded as

$$\begin{aligned}
\sum_{r=0}^{\infty} \frac{\beta_i^{2r+1}}{\prod_{n=0}^{r}(2n+1)} &= \beta_i[1 + \frac{1}{3}\beta_i^2 + \frac{1}{15}\beta_i^4 + \frac{1}{105}\beta_i^6 + \frac{1}{945}\beta_i^8 + ...\infty] && (21)\\
&\leq \beta_i[1 + \frac{1}{3}\beta_i^2 + \frac{1}{15}\beta_i^4 + \frac{1}{105}\beta_i^6 + \frac{1}{945}\beta_i^8 + ...\infty] && (22)\\
&= \beta_i \left[\frac{3}{3-\beta_i^2} \right], && (23)
\end{aligned}$$

where $\beta_i^2 < 3$.
Using the above series it can be simplified as

$$\frac{3\beta_i}{3-\beta_i^2} \geq \sqrt{\frac{\pi}{2}}(2y_i+1)\exp(\beta_i^2/2) \tag{24}$$

or,

$$\frac{3\beta_i}{3-\beta_i^2}\exp(-\beta_i^2/2) \geq \sqrt{\frac{\pi}{2}}(2y_i+1) \tag{25}$$

Hence the equivalent deterministic model of probabilistic problem (1)-(3) can be presented as

$$\begin{aligned}
\min Z^{(k)}(x) &= \sum_{j=1}^{n} c_j^{(k)} x_j,\ k=1,2,...,K && (26)\\
\text{subject to} &: \frac{3\beta_i}{3-\beta_i^2}\exp(-\beta_i^2/2) \geq \sqrt{\frac{\pi}{2}}(2y_i+1),\quad i=1,...,m, && (27)\\
\prod_{i=1}^{m} y_i &\geq 1-\alpha, && (28)\\
E(b_i) &= \sum_{i=1}^{n} a_{ij}x_j - \beta_i\sigma(b_i),\quad i=1,...,m, && (29)\\
& 0 \leq y_i \leq 1,\quad i=1,...,m, && (30)\\
& x_j \geq 0,\quad j=1,...,n. && (31)
\end{aligned}$$

3 Neural networks architectures for nonlinear mixed constraints

Let us consider the following deterministic multi-objective nonlinear programming problem:

$$\begin{aligned}
\min Z^{(k)}(x) &= \sum_{j=1}^{n} c_j^{(k)} x_j,\ k=1,2,...,K && (32)\\
\text{subject to} &: h_i(x)=0 \quad \text{where}\ \ i=1,2,...,p && (33)
\end{aligned}$$

and

$$g_i(x) \leq 0 \quad \text{where}\ \ i=p+1,p+2,...,m \tag{34}$$

where $x \in R^{n\times 1}$ is the vector of the independent variables,,$Z^{(k)}(x): R^{n\times m} \to R$ is the objective function, and functions $h_i(x), g_i(x): R^{n\times 1} \to R$ represent constraints. To simplify the derivations of the algorithms we will assume that both the objective function and the constraints are smooth differentiable functions of independent variables.

$$h_i(x) = \sum_{j=1}^{n} a_{ij}x_j - \beta_i\sigma(b_i) - E(b_i) = 0,\quad i=1,...,m \tag{35}$$

$$g_i(x) = \begin{pmatrix} (3-\beta_i^2)(2y_i+1)\sqrt{\frac{\pi}{2}} \leq 3\beta_i\exp(\frac{-\beta_i^2}{2}) \\ \prod_{i=1}^{m} y_i \leq \alpha-1 \\ 0 \leq y_i \leq 1,\quad i=1,...,m \\ x_j \geq 0,\quad j=1,...,n \end{pmatrix} \tag{36}$$

Using surplus variables, the inequality constraints can be converted to equality constraints. Similarly, each of the equality constraints can be converted to a pair of inequality constraints according to

$$h_i(x)=0 \Longleftrightarrow h_i(x)\leq 0 \quad \text{and} \quad h_i(x)\geq 0 \tag{37}$$

3.1 Neural Networks for Penalty Function for System of Nonlinear Equations

This Method using penalty functions make an attempt to transform the system of nonlinear equations (SNE) to an equivalent unconstrained optimization problem, or to a sequence of constrained optimization problems. This transformation is accomplished through modification of the objective function so that it includes terms

that penalize every violation of the constraints. In general, the modified objective function takes the following form:

$$f_A(x) = \sum_{i=1}^{p} K_i^{(1)} \Phi_i^{(1)}[h_i(x)] + \sum_{i=p+1}^{m} K_i^{(2)} \Phi_i^{(2)}[g_i(x)] \tag{38}$$

Functions $\Phi_i^{(1)}$ and $\Phi_i^{(2)}$ are called penalty functions, and they are designed to increase the value of the modified objective function $f_A(x)$ whenever the vector of independent variables violates a constraint, or in other words whenever it is outside the feasible region. Penalty functions are commonly selected as at least one-time differentiable function satisfying the following requirements:

1. For equality constraints

$$\Phi_i^{(1)} \begin{cases} > 0 & \text{for} \quad h_i(x) \neq 0 \\ = 0 & \text{for} \quad h_i(x) = 0 \end{cases} \tag{39}$$

2. For equality constraints

$$\Phi_i^{(2)} \begin{cases} > 0 & \text{for} \quad g_i(x) > 0 \\ = 0 & \text{for} \quad g_i(x) \leq 0 \end{cases} \tag{40}$$

For example, the typical modified objective function for the case3 problem can written as

$$f_A(x) = \sum_{i=1}^{p} \frac{K_i^{(1)}}{\rho_1} |h_i(x)|^{\rho_1} + \sum_{i=p+1}^{m} K_i^{(2)} \max\{0, g_i(x)\}^{\rho_2} \tag{41}$$

Where $\rho_1, \rho_2 \geq 0$. Parameters $K_i^{(1)}, K_i^{(2)} \geq 0$ are commonly referred to as penalty parameters or penalty multipliers, and in (41) we have assumed that a separate penalty parameter is associated with each of the penalty functions. In practice this is rarely the case, and commonly there is only one parameter mulitiplying the entire penalty term, that is,

$$f_A(x) = K \left[\sum_{i=1}^{p} \frac{1}{p_i} |h_i(x)|^{\rho_1} + \sum_{i=p+1}^{m} \max\{0, g_i(x)\}^{\rho_2} \right] = kp(x) \tag{42}$$

where $p(x)$ represents the penalty term.

There are two fundamental issues that need to be addressed in the practical application of penalty functions. First, we need to be aware that (42) represents merely an approximation of the original problem in (28) through (30). The question is. How close is the approximation? The second issue involves a design of a computationally efficient neural network algorithm that can successfully solve the unconstrained problem in a timely manner. From the form of the augmented objective function in (42), it should be obvious that the solution resides in the region where the value of the penalty function $P(x)$ is small. As a matter of fact, if K is increased toward infinity, the solution of the unconstrained problem will be forced into the feasible region of the original NP problem. Remember that if the point is in the feasible region, all the constraints are satisfied and the penalty function equals zero.

In the limiting case, when $k \to \infty$, the two problems become equivalent. Applying the steepest descent approach, we can generate the update equations in accordance with

$$x(k+1) = x(k) - \mu \frac{\partial f_A(x)}{\partial x} \tag{43}$$

where $\mu > 0$ is the learning rate parameter and the gradient term on the right hand side of (43) depends on the penalty function selection. For example, when the form of the energy function is as given in (42), with $\rho_1 = 2$ and $\rho_2 = 1$,we have

$$\frac{\partial f_A(x)}{\partial x} = K \left[\sum_{i=1}^{p} \frac{\partial h_i(x)}{\partial x} h_i(x) + \sum_{i=p+1}^{m} \frac{\partial}{\partial x} \max\{0, g_i(x)\} \right] \tag{44}$$

After substituting (44) into (43), we have for the learning rule

$$x(k+1) = x(k) - \mu \left[K \sum_{i=1}^{p} \frac{\partial h_i(x)}{\partial x} h_i(x) + K \sum_{i=p+1}^{m} \frac{\partial}{\partial x} \max\{0, g_i(x)\} \right] \tag{45}$$

The neural network architecture realization of this process is presented in figure 1 Note that only a portion of the network for computing a single component of the independent-variable vector is presented. Also note that the network corresponds to the general case of this problem.

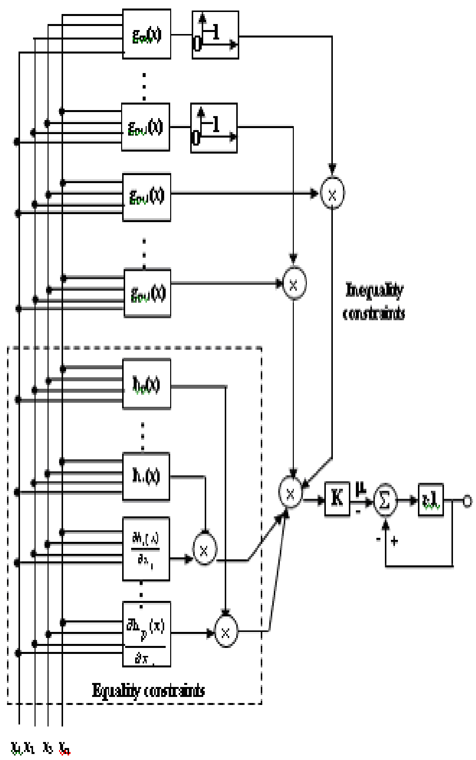


Figure 1: Discrete-tim network for deterministic nonlinear constraints implementing penalty function method, implementation of Equation (45).

4 Neural Networks For Nonlinear Continuous Constrained Optimization Problems

The (DNP) problem with mixed constraints has the following general form:

$$\text{Minimize} \quad \lambda \tag{46}$$

$$\text{Subject to:} \xi_i(x) = \sum_{j=1}^{n} a_{ij} x_j - \beta_i \sigma(b_i) - E(b_i) = 0, \quad i = 1, ..., m, \tag{47}$$

and

$$\psi_i(x) = \begin{pmatrix} c_{ij} x_j - \lambda_i \le Z_i \\ (3 - \beta_i^2)(2y_i + 1)\sqrt{\frac{\pi}{2}} \le 3\beta_i \exp(\frac{-\beta_i^2}{2}) \\ \prod_{i=1}^{m} y_i \le \alpha - 1 \\ 0 \le y_i \le 1, \quad i = 1, ..., m \\ x_j \ge 0, \quad j = 1, ..., n \end{pmatrix} \tag{48}$$

Using surplus variables, the inequality constraints can be converted to equality constraints. Similarly, each of the equality constraints can be converted to a pair of inequality constraints according to

$$\xi_i(x) = 0 \leftrightarrow \xi_i(x) \leq 0 \ \text{ and } \ \xi_i(x) \geq 0 \tag{49}$$

4.1 Neural Networks for Penalty Function DNP Methods

The modified objective function for this problem can written as

$$R_A(x) = \lambda + \sum_{i=1}^{p} \frac{K_i^{(1)}}{\rho_1} |\xi_i(x)|^{\rho_1} + \sum_{i=p+1}^{m} K_i^{(2)} \max\{0, \psi_i(x)\}^{\rho_2} \tag{50}$$

In practice this is rarely the case, and commonly there is only one parameter multiplying the entire penalty term, that is,

$$R_A(x) = \lambda + K \left[\sum_{i=1}^{p} \frac{1}{\rho_1} |\xi_i(x)|^{\rho_1} + \sum_{i=p+1}^{m} \max\{0, \psi_i(x)\}^{\rho_2} \right] = \lambda + kp(x) \tag{51}$$

where $p(x)$ represents the penalty term.

Applying the steepest descent approach, we can generate the update equations in accordance with

$$x(k+1) = x(k) - \mu \frac{\partial R_A(x)}{\partial x} \tag{52}$$

where $\mu > 0$ is the learning rate parameter and the gradient term on the right hand side of (52) depends on the penalty function selection. For example, when the form of the energy function is as given in (51), with $\rho_1 = 2$ and $\rho_2 = 1$,we have

$$\frac{\partial R_A(x)}{\partial x} = \lambda + K \left[\sum_{i=1}^{p} \frac{\partial \xi_i(x)}{\partial x} \xi_i(x) + \sum_{i=p+1}^{m} \frac{\partial}{\partial x} \max\{0, \psi_i(x)\} \right] \tag{53}$$

After substituting (53) into (52), we have for the learning rule

$$x(k+1) = x(k) - \mu \left[\lambda + K \sum_{i=1}^{p} \frac{\partial \xi_i(x)}{\partial x} \xi_i(x) + K \sum_{i=p+1}^{m} \frac{\partial}{\partial x} \max\{0, \psi_i(x)\} \right] \tag{54}$$

The neural network architecture realization of this process is presented in figure 2.

5 Numerical example

$$\begin{aligned}
\text{min :} \quad & 2x_1 + 3x_2 + 6x_3. \\
\text{min :} \quad & x_1 + 2x_2 + 4x_3 \\
\text{subject to:} \quad & Pro[2x_1 + x_2 + 2x_3 \geq b_1, \\
& x_1 + 2x_2 + 4x_3 \geq b_2] \geq 8.85, \\
x_1, x_2, x_3 \geq 0. &
\end{aligned}$$

The mean and standard deviation of b_1, and b_2 are give as

$$E(b_1) = 6. \quad \sigma(b_1) = 2$$

$$E(b_2) = 7. \quad \sigma(b_2) = 4.$$

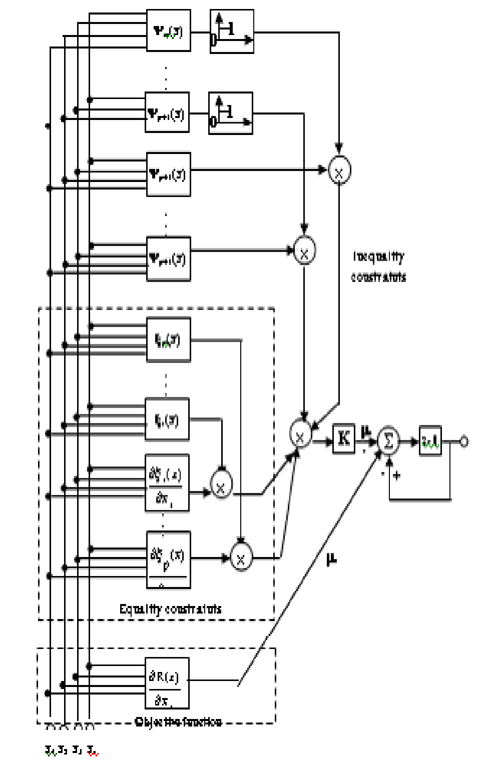


Figure 2: Discrete-tim network for deterministic nonlinear constraints implementing penalty function method, implementation of Equation (45).

We obtain the equivalent deterministic programming problem for the above multi-objective stochastic programming problem by using Eqs. (20)-(25).

$$\begin{aligned}
\min: \quad & z^{(1)} = 2x_1 + 3x_2 + 6x_3. \\
\min: \quad & z^{(2)} = x_1 + 2x_2 + 4x_3. \\
\text{subject to:} \quad & 2x_1 + x_2 + 2x_3 - 3\beta_1 = 6. \\
& x_1 + 2x_2 + 4x_3 - 4\beta_1 = 7. \\
& 1.2533141(1+2y_1)(3-\beta_1^2) \le 3\beta_1 \exp(\frac{-\beta_1^2}{2}), \\
& 1.2533141(1+2y_2)(3-\beta_2^2) \le 3\beta_2 \exp(\frac{-\beta_2^2}{2}), \\
& y_1 y_2 \ge 0.85, \\
& y_1, y_2 \le 1, \\
& x_1, x_2, x_3, y_1, y_2 \ge 0, \\
& \text{ and } \beta_1 \text{ and } \beta_2 \text{ are unrestricted in sign.}
\end{aligned}$$

The above problem is obviously an NP problem with inequality constraints. The modified penalty function can be formed as:
$F_A(x) = (3x_1+3x_2+6x_3-3\beta_1-4\beta_2-13)+k[0, 1.2(\beta_1^2+\beta_2^2)-7.6(y_1+y_2)+2.4(y_1\beta_1^2+y_2\beta_2^2)+3(\beta_1\exp(-\beta_1^2/2)+\beta_2\exp(-\beta_2^2/2)+y_1y_2-8.05)]$
By using the steepest descent method, the update equations can be computed as

$$\begin{aligned}
x_1(k+1) &= x_1(k+1) - \mu k(5x_1 + 4x_2 + 8x_3 - 6\beta_1 - 4\beta_2 - 19) \\
x_2(k+1) &= x_2(k+1) - \mu k(4x_1 + 5x_2 + 16x_3 - 3\beta_1 - 8\beta_2 - 20) \\
x_3(k+1) &= x_3(k+1) - \mu k(8x_1 + 10x_2 + 20x_3 - 6\beta_1 - 16\beta_2 - 40) \\
\beta_1(k+1) &= \beta_{1(k)} - \mu k(2.4\beta_1(1+2y_1) + 3(1 - \beta_1^2\exp(-\beta_1^2/2)) \\
\beta_2(k+1) &= \beta_{2(k)} - \mu k(2.4\beta_2(1+2y_2) + 3(1 - \beta_2^2\exp(-\beta_2^2/2)) \\
y_1(k+1) &= y_1(k) - \mu k(-7.6 + 2.4\beta_1^2 + y_2) \\
y_2(k+1) &= y_2(k) - \mu k(-7.6 + 2.4\beta_2^2 + y_1)
\end{aligned}$$

The neural network architecture shown in figure 1 was used to determine the solution of the NP problem. Parameters of the network were chosen as $k = 1$, and $\mu = 0.001$, and initial solution was set as $x = [2\ \ 1\ \ 1.5\ \ 1\ \ 1\ \ 0.5\ \ 0.5]$. The network converged in approximately 300 iterations, and the optimal solution given by.

$$x^{(1)} = \begin{pmatrix} 2.770882 \\ 1.680341 \\ 1.833226 \end{pmatrix}, \quad \beta^{(1)} = \begin{pmatrix} 1.629519 \\ 1.616117 \end{pmatrix}, \quad y^{(1)} = \begin{pmatrix} 1.0000 \\ 0.8500 \end{pmatrix}$$

Substitute by this values , we formulate the following nonlinear programming problem:

$$\begin{aligned}
\max: \quad & \lambda \\
\text{subject to:} \quad & 2x_1 + 3x_2 + 6x_3 + \lambda = 21.582143, \\
& x_1 + 2x_2 + 4x_3 + \lambda = 13.464468, \\
& 2x_1 + x_2 + 2x_3 - 3\beta_1 = 6, \\
& x_1 + 2x_2 + 4x_3 - 4\beta_2 = 7, \\
& 1.2533141(1+2y_1)(3-\beta_1^2) \le 3\beta_1 exp(-\beta_1^2/2), \\
& 1.2533141(1+2y_2)(3-\beta_2^2) \le 3\beta_2 exp(-\beta_2^2/2), \\
& y_1 y_2 \ge 0.85, \\
& y_1, y_2 \le 1, \\
& x_1, x_2, x_3, y_1, y_2 \ge 0, \\
& \text{and}\beta_1\text{and}\beta_2\text{are unrestricted in sign.}
\end{aligned}$$

The penalty function can be formed as:
$R_A(x) = \lambda + (3x_1 + 3x_2 + 6x_3 - 3\beta_1 - 4\beta_2 - 13) + k[0, 1.2(\beta_1^2 + \beta_2^2) - 6(y_1 + y_2) + 2.4(y_1\beta_1^2 + y_2\beta_2^2) + 3(\beta_1 \exp(-\beta_1^2/2) + \beta_2 \exp(-\beta_2^2/2) + y_1 y_2 - 8.05)]$
And solving by using the steepest descent method, the update equations can be computed as

$$\begin{aligned}
x_1^*(k+1) &= x_1^*(k) - \mu k(10x_1 + 12x_2 + 24x_3 - 6\beta_1 - 4\beta_2 + 3\lambda - 37.626619) \\
x_2^*(k+1) &= x_2^*(k+1) - \mu k(12x_1 + 18x_2 + 42x_3 - 3\beta_1 - 8\beta_2 + 5\lambda - 54.6688) \\
x_3^*(k+1) &= x_3^*(k+1) - \mu k(24x_1 + 36x_2 + 72x_3 - 6\beta_1 - 16\beta_2 + 10\lambda - 109.3377) \\
\lambda(k+1) &= \lambda(k) - \mu k(3x_1 + 5x_2 + 10x_3 + 2\lambda - 16.0444) \\
\beta_1^*(k+1) &= \beta_{1(k)}^* - \mu k(2.4\beta_1(1 + 2y_1) + 3(1 - \beta_1^2 exp(-\beta_1^2/2)) \\
\beta_2^*(k+1) &= \beta_{2(k)}^* - \mu k(2.4\beta_2(1 + 2y_1) + 3(1 - \beta_2^2 exp(-\beta_2^2/2)) \\
y_1^*(k+1) &= y_1^*(k) - \mu k(-7.6 + 2.4\beta_1^2 + y_2) \\
y_2^*(k+1) &= y_2^*(k) - \mu k(-7.6 + 2.4\beta_2^2 + y_1)
\end{aligned}$$

To solve this SMONLP problem, the neural network in Figure 2 is simulated. The parameters of the network were chosen as $\mu = 0.01$, and $k = 1$. The network converges in approximately 100 iterations. The optimal solution to the SMONLP is given as

$$x^{*(1)} = \begin{pmatrix} 2.770882 \\ 1.678597 \\ 1.834098 \end{pmatrix}, \quad \beta^{*(1)} = \begin{pmatrix} 1.629519 \\ 1.616117 \end{pmatrix}, \quad y^{*(1)} = \begin{pmatrix} 1.0000 \\ 0.849998 \end{pmatrix} \quad \text{and} \quad \lambda = 0.987653$$

6 Conclusions

We have proposed a recurrent neural network for solving stochastic multi-objective programming problems with general nonlinear constraints. The proposed neural network has a simpler structure and a lower complexity for implementation than the existing neural networks for solving such problems. It is shown here that the proposed neural network is stable in the sense of Lyapunov and globally convergent to an optimal solution. Compared with the existing convergence results, the present results do not require Lipschitz continuity condition on the stochastic multi-objective objective function. Finally, examples are provided to show the performance of the proposed neural network.

References

[1] A. Charnes and W. Cooper, " Chance constrained programming ", Management Sci. 6 (1959) 73-79.

[2] Fredric M. Ham " Principles of Neurocomputing for Science and Engineering " McGraw-Hill Higher Education, International Edition 2001

[3] H. Li, B. S. Manjunath, and S. K. Mitra, "Multisensor image fusion using the wavelet transform," Graph. Models Image Process., vol. 57, no. 3, pp. 235-245, 1995.

[4] J.P.Leclercq, "Stochastic programming: an interactive multicriteria approach", Eur. J. Oper. Res. 10 (1982)33-41.

[5] J. Jr. Teghem, D. Dufrance, M. Thauvoye and P. Kunch, Strange, "An interactive method for multi-objective linear programming under uncertainty", Eur. J. Oper. Res. 26 (1986) 65-82.

[6] M.P. Kennedy and L. O. Chua, "Neural networks for nonlinear programming," IEEE Transactions on Circuits and Systems, vol. 35, no. 5, 554-562, 1988.

[7] M. El-Sayed Wahed " Neural Network Representation of Stochastic Quadratic Programming Problems with bi's and cj's Follow Goint Distribution" Jour. Inst. Maths. & Computer Sciences, Vol1. 15, No.2, pp. 263-279, (2004).

[8] M. El-Sayed Wahed " Neural Network For Stochastic Linear Programming Problems " Egyption Informatics Journal , Vol. 5/ No.2, pp. 28-42, (2004).

[9] M. R. Hestens, "Multiplier Gradient Methods," Journal of Optimization Theory and Applications, vol. 4, pp. 303-320, (1969).

[10] M. Wahed, "On some solution for solving a stochastic programming problem with known probability distribution", M. Sc. Thesis, Department of Computer Science, Institute of Statistical Studies and Research, Cairo University, (1991).

[11] N.S. Kambo, "Mathematical Programming Techniques", (Affiliated East-West Press Pvt. Ltd., 1984).

[12] Q. Tao, J. D. Cao, M. S. Xue, and H. Qiao, "A high performance neural network for solving nonlinear programming problems with hybrid constraints," Phys. Lett. A, vol. 288, no. 2, pp. 88-94, 2001.

[13] S. S. Rao, "Optimization Theory and Applications", (Wiley Eastern Limited, New Delhi, 2nd edn,1984).

[14] W. Mohamed, "Fuzzy programming approach to multi-objective stochastic linear programming problems with a known distribution", The Journal of Fuzzy Mathematics Vol. 7, No. 3, (1999)

[15] W. Mohammed, "Chance constrained fuzzy goal programming with right-hand side uniform random variable coefficients", Fuzzy sets and systems 109 (2000) 107-110.

[16] W.H. Press, B.P. Flannery, S.A. Teukolsky, W.T. Vetterling, Numerical Recipes in C, Cambridge University Press, Cambridge.

[17] Y.S. Xia and J. Wang, "A general methodology for designing globally convergent optimization neural networks," IEEE Trans. Neural Netw., vol. 9, no. 6, pp. 1331-1343, Nov. 1998.

[18] Y.S. Xia and G. Feng "A Modified Neural Network for Quadratic Programming with Real-Time Applications", Neural Information Processing-Letters and Reviews, vol. 3, no.3, pp. 69-76, 2004.

JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS,VOL.5,NO.3,231-236,2007,COPYRIGHT 2007 EUDOXUS PRESS,LLC

Absolute Cesàro Summability Factors

Hikmet Seyhan Özarslan

Department of Mathematics, Erciyes University, 38039 Kayseri, Turkey

E-mail:seyhan@erciyes.edu.tr

Abstract

In this paper a general theorem concerning the $\varphi - \mid C, 1 \mid_k$ summability factors of infinite series, has been proved.

2000 AMS Subject Classification: 40F05, 40D15, 40D25

Keywords: Absolute Cesàro summability, infinite series, summability factors.

1 Introduction

Let (φ_n) be a sequence of complex numbers and let $\sum a_n$ be a given infinite series with partial sums (s_n). We denote by z_n^α and t_n^α the n-th Cesàro means of order α, with $\alpha > -1$, of the sequence (s_n) and (na_n), respectively, i.e.,

$$z_n^\alpha = \frac{1}{A_n^\alpha} \sum_{v=0}^{n} A_{n-v}^{\alpha-1} s_v, \tag{1}$$

$$t_n^\alpha = \frac{1}{A_n^\alpha} \sum_{v=1}^{n} A_{n-v}^{\alpha-1} v a_v, \tag{2}$$

where

$$A_n^\alpha = O(n^\alpha), \quad \alpha > -1, \quad A_0^\alpha = 1 \quad and \quad A_{-n}^\alpha = 0 \quad for \quad n > 0. \tag{3}$$

The series $\sum a_n$ is said to be summable $\varphi - \mid C, \alpha \mid_k$, $k \geq 1$ and $\alpha > -1$, if (see [2])

$$\sum_{n=1}^{\infty} \mid \varphi_n (z_n^\alpha - z_{n-1}^\alpha) \mid^k < \infty. \tag{4}$$

But since $t_n^\alpha = n(z_n^\alpha - z_{n-1}^\alpha)$ (see [3]) condition (4) can also be written as

$$\sum_{n=1}^{\infty} n^{-k} \mid \varphi_n t_n^\alpha \mid^k < \infty. \tag{5}$$

In the special case when $\varphi_n = n^{1-\frac{1}{k}}$ (resp. $\varphi_n = n^{\delta+1-\frac{1}{k}}$) $\varphi - \mid C, \alpha \mid_k$ summability is the same as $\mid C, \alpha \mid_k$ (resp. $\mid C, \alpha; \delta \mid_k$) summability. Also if we take $\alpha = 1$, $\varphi - \mid C, \alpha \mid_k$ summability reduces to $\varphi - \mid C, 1 \mid_k$ summability.

The series $\sum a_n$ is said to be bounded $[R, logn, 1]_k$, $k \geq 1$, if (see [6])

$$\sum_{v=1}^{n} \frac{\mid s_v \mid^k}{v} = O(logn) \quad as \quad n \to \infty. \tag{6}$$

A sequence (λ_n) is said to be convex if $\Delta^2 \lambda_n \geq 0$, for $n = 1, 2, ...$ where $\Delta\lambda_n = \lambda_n - \lambda_{n+1}$. Mishra [6] (see also Mazhar [4]) proved the following theorem.

Theorem A. Let (λ_n) be a convex sequence such that $\sum \frac{\lambda_n}{n}$ is convergent. If $\sum a_n$ is bounded $[R, logn, 1]_k$, then $\sum a_n \lambda_n$ is summable $\mid C, 1 \mid_k$, $k \geq 1$.

Mishra and Srivastava [7] have proved Theorem A under very general and weaker conditions in the following form.

Theorem B. Let (X_n) be a positive non-decreasing sequence and there be sequences (β_n) and (λ_n) such that

$$\mid \Delta\lambda_n \mid \leq \beta_n, \tag{7}$$

$$\beta_n \to 0 \quad as \quad n \to \infty, \tag{8}$$

$$\sum_{n=1}^{\infty} n \mid \Delta\beta_n \mid X_n < \infty, \tag{9}$$

$$\mid \lambda_n \mid X_n = O(1) \quad as \quad n \to \infty. \tag{10}$$

If

$$\sum_{v=1}^{n} \frac{\mid s_v \mid^k}{v} = O(X_n) \quad as \quad n \to \infty, \tag{11}$$

then the series $\sum a_n \lambda_n$ is summable $\mid C, 1 \mid_k$, $k \geq 1$.

2. The main result. The aim of this paper is to generalize Theorem B under weaker conditions for $\varphi - \mid C, 1 \mid_k$ summability. For this we need the concept of almost increasing

sequence. A positive sequence (b_n) is said to be almost increasing if there exists a positive increasing sequence (c_n) and two positive constants A and B such that $Ac_n \leq b_n \leq Bc_n$ (see [1]). Obviously every increasing sequence is an almost increasing sequence, but the converse need not be true, as can be seen from the example $b_n = ne^{(-1)^n}$. So we are weakening the hypotheses of the theorem in replacing the increasing sequence by an almost increasing sequence.

Now, we shall prove the following theorem:

Theorem. Let (X_n) be an almost increasing sequence and the sequences (β_n) and (λ_n) such that conditions (7)-(10) of Theorem B are satisfied. If there exists an $\epsilon > 0$ such that the sequence $(n^{\epsilon-k} \mid \varphi_n \mid^k)$ is non-increasing and

$$\sum_{v=1}^{n} v^{-k} \mid \varphi_v s_v \mid^k = O(X_n) \quad as \quad n \to \infty, \tag{12}$$

then the series $\sum a_n\lambda_n$ is summable $\varphi- \mid C, 1 \mid_k$, $k \geq 1$.

We need the following lemma for the proof of our theorem.

Lemma 1 ([5]). Under the conditions on (X_n), (β_n) and (λ_n) as taken in the statement of the theorem, the following conditions hold, when (9) is satisfied:

$$\sum_{n=1}^{\infty} \beta_n X_n < \infty \tag{13}$$

and

$$n\beta_n X_n = O(1) \quad as \quad n \to \infty. \tag{14}$$

3. Proof of the Theorem. Let (u_n) and (T_n) be the n-th $(C, 1)$ means of the series $\sum a_n$ and the sequence (na_n), respectively. Now,we will show that

$$\sum_{n=1}^{\infty} \frac{1}{n^k} \mid \varphi_n T_n \mid^k < \infty, \tag{15}$$

where

$$T_n = \frac{1}{n+1} \sum_{v=1}^{n} v a_v \lambda_v. \tag{16}$$

Now, applying Abel's transformation to the sum, we get

$$\begin{aligned} T_n &= \frac{1}{n+1} \sum_{v=1}^{n-1} v \Delta\lambda_v s_v - \frac{1}{n+1} \sum_{v=1}^{n-1} \lambda_{v+1} s_v + \frac{n s_n \lambda_n}{n+1} - \frac{a_0 \lambda_1}{n+1} \\ &= T_{n,1} + T_{n,2} + T_{n,3} + T_{n,4}, \quad say. \end{aligned}$$

Since

$$| T_{n,1} + T_{n,2} + T_{n,3} + T_{n,4} |^k \leq 4^k (| T_{n,1} |^k + | T_{n,2} |^k + | T_{n,3} |^k + | T_{n,4} |^k),$$

to complete the proof of the theorem, it is sufficient to show that

$$\sum_{n=1}^{\infty} \frac{1}{n^k} | \varphi_n T_{n,r} |^k < \infty \quad for \quad r = 1,2,3,4. \tag{17}$$

Now, when $k > 1$, applying Hölder's inequality with indices k and k', where $\frac{1}{k} + \frac{1}{k'} = 1$ and later using Abel's transformation, we get

$$\begin{aligned}
\sum_{n=2}^{m+1} \frac{1}{n^k} | \varphi_n T_{n,1} |^k &\leq \sum_{n=2}^{m+1} \frac{| \varphi_n |^k}{n^{k+1}} \{ \sum_{v=1}^{n-1} v | \Delta\lambda_v || s_v |^k \} \times \{ \frac{1}{n} \sum_{v=1}^{n-1} v | \Delta\lambda_v | \}^{k-1} \\
&= O(1) \sum_{v=1}^{m} v | \Delta\lambda_v || s_v |^k \sum_{n=v+1}^{m+1} \frac{| \varphi_n |^k}{n^{k+1}} \\
&= O(1) \sum_{v=1}^{m} v | \Delta\lambda_v || s_v |^k | \varphi_v |^k v^{\epsilon-k} \sum_{n=v+1}^{m+1} \frac{1}{n^{\epsilon+1}} \\
&= O(1) \sum_{v=1}^{m} v | \Delta\lambda_v | | \varphi_v s_v |^k v^{\epsilon-k} \int_v^{\infty} \frac{dx}{x^{\epsilon+1}} \\
&= O(1) \sum_{v=1}^{m} v | \Delta\lambda_v | v^{-k} | \varphi_v s_v |^k \\
&= O(1) \sum_{v=1}^{m} v\beta_v v^{-k} | \varphi_v s_v |^k \\
&= O(1) \sum_{v=1}^{m-1} \Delta(v\beta_v) \sum_{r=1}^{v} r^{-k} | \varphi_r s_r |^k + O(1) m\beta_m \sum_{v=1}^{m} v^{-k} | \varphi_v s_v |^k \\
&= O(1) \sum_{v=1}^{m-1} | \Delta(v\beta_v) | X_v + O(1) m\beta_m X_m \\
&= O(1) \sum_{v=1}^{m-1} | (v+1)\Delta\beta_v - \beta_v | X_v + O(1) m\beta_m X_m \\
&= O(1) \sum_{v=1}^{m-1} v | \Delta\beta_v | X_v + O(1) \sum_{v=1}^{m-1} | \beta_v | X_v + O(1) m\beta_m X_m \\
&= O(1) \quad as \quad m \to \infty,
\end{aligned}$$

by virtue of the hypotheses of the Theorem and Lemma.

Again, since (λ_n) is bounded, as in $T_{n,1}$, we have that

$$
\begin{aligned}
\sum_{n=2}^{m+1} \frac{1}{n^k} \mid \varphi_n T_{n,2} \mid^k &\leq \sum_{n=2}^{m+1} \frac{\mid \varphi_n \mid^k}{n^{k+1}} \{\sum_{v=1}^{n-1} \mid \lambda_{v+1} \mid^k \mid s_v \mid^k\} \times \{\frac{1}{n} \sum_{v=1}^{n-1} 1\}^{k-1} \\
&= O(1) \sum_{v=1}^{m} \mid \lambda_{v+1} \mid^k \mid s_v \mid^k \sum_{n=v+1}^{m+1} \frac{\mid \varphi_n \mid^k}{n^{k+1}} \\
&= O(1) \sum_{v=1}^{m} \mid \lambda_{v+1} \mid^k v^{-k} \mid \varphi_v s_v \mid^k \\
&= O(1) \sum_{v=1}^{m} \mid \lambda_{v+1} \mid^{k-1} \mid \lambda_{v+1} \mid v^{-k} \mid \varphi_v s_v \mid^k \\
&= O(1) \sum_{v=1}^{m} \mid \lambda_v \mid v^{-k} \mid \varphi_v s_v \mid^k \\
&= O(1) \sum_{v=1}^{m-1} \Delta \mid \lambda_v \mid \sum_{r=1}^{v} r^{-k} \mid \varphi_r s_r \mid^k + O(1) \mid \lambda_m \mid \sum_{v=1}^{m} v^{-k} \mid \varphi_v s_v \mid^k \\
&= O(1) \sum_{v=1}^{m-1} \mid \Delta\lambda_v \mid X_v + O(1) \mid \lambda_m \mid X_m \\
&= O(1) \sum_{v=1}^{m-1} \beta_v X_v + O(1) \mid \lambda_m \mid X_m \\
&= O(1) \quad as \quad m \to \infty,
\end{aligned}
$$

by virtue of the hypotheses of the Theorem and Lemma.

Also, as in $T_{n,2}$, we have

$$
\begin{aligned}
\sum_{n=1}^{m} \frac{1}{n^k} \mid \varphi_n T_{n,3} \mid^k &= O(1) \sum_{n=1}^{m} \mid \lambda_n \mid n^{-k} \mid \varphi_n s_n \mid^k \\
&= O(1) \quad as \quad m \to \infty.
\end{aligned}
$$

Finally, since $(n^{\epsilon-k} \mid \varphi_n \mid^k)$ is non-increasing, by hypothesis, we have

$$
\begin{aligned}
\sum_{n=1}^{m} \frac{1}{n^k} \mid \varphi_n T_{n,4} \mid^k &= O(1) \sum_{n=1}^{m} \frac{\mid \varphi_n \mid^k}{n^{2k}} \\
&= O(1) \sum_{n=1}^{m} \frac{n^{\epsilon-k} \mid \varphi_n \mid^k}{n^{\epsilon+k}} \\
&= O(1) \sum_{n=1}^{m} \frac{1}{n^{\epsilon+k}} \\
&= O(1) \quad as \quad m \to \infty, (\epsilon > 0, k \geq 1).
\end{aligned}
$$

Therefore, we get

$$\sum_{n=1}^{m} n^{-k} \mid \varphi_n T_{n,r} \mid^k = O(1) \quad as \quad m \to \infty, \quad for \quad r = 1, 2, 3, 4.$$

This completes the proof of the Theorem.

4. Special cases:

1. If we take (λ_n) as a convex sequence such that $\sum n^{-1}\lambda_n$ is convergent, $X_n = logn, \epsilon = 1$ and $\varphi_n = n^{1-k^{-1}}$ in our theorem, then we get Theorem A.

2. If we take (X_n) as a positive non-decreasing sequence, $\epsilon = 1$ and $\varphi_n = n^{1-k^{-1}}$ in our theorem, then we get Theorem B.

References

[1] S. Aljancic and D. Arandelovic, O-regularly varying functions, Publ. Inst. Math., 22 (1977), 5-22.

[2] M. Balcı, Absolute φ-summability factors, Comm. Fac. Sci. Univ. Ankara, Ser. A_1 29 (1980), 63-80.

[3] E. Kogbetliantz, Sur lés series absolument sommables par la méthode des moyennes arithmétiques, Bull. Sci. Math., 49 (1925), 234-256.

[4] S.M. Mazhar, On $\mid C, 1 \mid_k$ summability factors of infinite series, Acta Sci. Math. Szeged, 27 (1966), 67-70.

[5] S.M. Mazhar, A note on absolute summability factors, Bull. Inst. Math. Acad. Sinica, 25 (1997), 233-242.

[6] B.P. Mishra, On absolute Cesàro summability factors of infinite series, Rend. Circl. Mat. Palermo, (2), 14 (1965), 189-193.

[7] K.N. Mishra and R.S.L. Srivastava, On absolute Cesàro summability factors of infinite series, Portugaliae Math., 42 (1983-1984), 53-61.

JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS,VOL.5,NO.3,237-254,2007,COPYRIGHT 2007 EUDOXUS PRESS,LLC

Analytical and Numerical Solutions of the Inhomogeneous Heat Equation

T. Matsuura and S. Saitoh
Department of Mechanical Engineering,
Faculty of Engineering,
Gunma University, Kiryu 376-8515, Japan
Department of Mathematics,
Faculty of Engineering,
Gunma University, Kiryu 376-8515, Japan
E-mail: matsuura@me.gunma-u.ac.jp,
ssaitoh@math.sci.gunma-u.ac.jp

Abstract

In this paper, by a new concept and method we shall give practical and numerical solutions of the inhomogeous heat equation on multi-dimensional spaces and show their numerical experiments by using computers.

Keywords: Heat equation, approximation of functions, reproducing kernel, Tikhonov regularization, Sobolev space, generalized inverse, approximate inverse, error estimate, noise, weighted convolution inequality

Mathematics Subject Classification (2000): Primary 44A15;35K05;30C40

1 INTRODUCTION AND MAIN RESULTS

We shall give simple approximate solutions for the inhomogeous heat equation, for any $L_2(\mathbf{R}^n)$ function g

$$Hu(x,t) = \partial_t u(x,t) - \Delta_x u(x,t) = g \quad \text{on} \quad \mathbf{R}^n \tag{1.1}$$

in the class of the functions of the s order Sobolev Hilbert space H^s on the whole real space $\mathbf{R}^n (n \geq 2, s \geq 2, s > n/2)$. In this paper, we shall use the notation

$$x' = (x_1, x_2, ..., x_{n-1}), x_n = t$$

and similarly

$$p = (p', p_n) \in \mathbf{R}^n.$$

This equation is, of course, very fundamental and has many applications to mathematical sciences.

Recently, in those papers [1,5,6], we were able to obtain very and surprisingly simple and practical approximate real inversion formulas for the very difficult Gaussian convolution equation by using the theory of reproducing kernels from the ideas of best approximations and generalized inverses. Furthermore, we illustrated their numerical experiments by using computers and we can realize that we were able to obtain practical real inversion formulas in [1]. Incidentally, their ideas were naturally combined with the idea and the method of the Tikhonov regularization by using the theory of reproducing kernels.

In this paper, we shall examine the corresponding problems for the heat equation on multidimensional spaces. We can obtain similar formulas and so, we are, in particular, interested in their numerical experiments by using computers. Furthermore, we shall establish error estimates for our solutions, because practical data contain errors and noises.

We shal recall the m order Sobolev Hilbert space H^m comprising functions F on $\mathbf{R}^n$ with the norm

$$\|F\|_{H^m}^2$$

$$= \sum_{\nu=0}^{m} {}_mC_\nu \sum_{r_1, r_2, \ldots, r_n \geq 0}^{\nu} \frac{\nu!}{r_1! r_2! \cdots r_n!} \int_{\mathbf{R}^n} \left(\frac{\partial^\nu F(x)}{\partial x_1^{r_1} \partial x_2^{r_2} \cdots \partial x_n^{r_n}} \right)^2 dx. \tag{1.2}$$

Here, of course,

$$r_1 + r_2 + \cdots + r_n = \nu.$$

This Hilbert space admits the reproducing kernel

$$K(x,y) = \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} \frac{1}{(1 + |\xi|^2)^m} e^{i(x-y)\cdot\xi} d\xi \tag{1.3}$$

as we see easily by using Fourier's transform (cf. [3], page 58). Note that the Sobolev Hilbert space H^s admitting the reproducing kernel (1.3) for $m = s$ can be defined for any positive number $s (s > n/2)$ in terms of Fourier integrals $\hat{F}$ of F

$$\hat{F}(\xi) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbf{R}^n} e^{-i\xi \cdot x} F(x) dx$$

as follows:

$$\|F\|_{H^s}^2 = \int_{\mathbf{R}^n} |\hat{F}(\xi)|^2 (1 + |\xi|^2)^s d\xi.$$

Our formulations and results are stated as follows:

THEOREM 1 *Let $n \geq 2, s \geq 2$ and $s > n/2$. For any function $g \in L_2(\mathbf{R}^n)$ and for any $\lambda > 0$, the best approximate function $F^*_{\lambda,s,g}$ in the sense*

$$\inf_{F \in H^s} \left\{ \lambda \|F\|_{H^s}^2 + \|g - HF\|_{L_2(\mathbf{R}^n)}^2 \right\}$$

$$= \lambda \|F^*_{\lambda,s,g}\|_{H^s}^2 + \|g - HF^*_{\lambda,s,g}\|_{L_2(\mathbf{R}^n)}^2 \tag{1.4}$$

exists uniquely and $F^*_{\lambda,s,g}$ is represented by

$$F^*_{\lambda,s,g}(x) = \int_{\mathbf{R}^n} g(\xi) Q_{\lambda,s}(\xi - x) d\xi \tag{1.5}$$

for

$$Q_{\lambda,s}(\xi - x) = \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} \frac{(-ip_n + |p'|^2) e^{-ip\cdot(\xi - x)} dp}{\lambda(|p|^2 + 1)^s + (p_n^2 + |p'|^4)}.$$

If, for $F \in H^s$ we consider the function $(HF)(x)$ and we take it as g, then we have the (favourable) result: as $\lambda \to 0$

$$F^*_{\lambda,s,g} \to F, \tag{1.6}$$

uniformly.

When the practical data g contain errors or noises, we need error estimates for the approximate solution (1.5). Following the idea of weighted convolution inequalities in [4], we obtain

THEOREM 2 *In the representation of the approximate solution (1.5), we obtain the estimate*

$$\int_{\mathbf{R}^n} |F^*_{\lambda,s,g}(x)|^2 dx \leq \frac{\Gamma(s-n/2)}{2^{n+2}\Gamma(s)} \frac{1}{\lambda} \int_{\mathbf{R}^n} |g(\xi)|^2 e^{|\xi|^2} d\xi. \tag{1.7}$$

The estimate (1.7) implies that for g_δ containing errors and noises,

$$\int_{\mathbf{R}^n} |F^*_{\lambda,s,g}(x) - F^*_{\lambda,s,g_\delta}(x)|^2 dx \leq \frac{\Gamma(s-n/2)}{2^{n+2}\Gamma(s)} \frac{1}{\lambda} \int_{\mathbf{R}^n} |g(\xi) - g_\delta(\xi)|^2 e^{|\xi|^2} d\xi.$$

Meanwhile, in Theorem 1 we wish to take a small λ in order to obtain a true solution F. Therefore, for

$$\int_{\mathbf{R}^n} |g(\xi) - g_\delta(\xi)|^2 e^{|\xi|^2} d\xi \leq \delta,$$

we wish to take δ and λ as follows:

$$\delta \to 0$$

and

$$\frac{\delta}{\lambda} \to 0.$$

The integral weight $e^{|\xi|^2}$ will be acceptable, because the functions g and g_δ decay exponentially or we can assume that they have compact supports.

In terms of the Sobolev norm, we obtain

THEOREM 3 *Let $\delta > 0$ and let g, g_δ satisfy*

$$\|g - g_\delta\|_{L_2(\mathbf{R}^n)} \leq \delta.$$

Then, we have

$$\|F^*_{\lambda,s,g_\delta} - F^*_{\lambda,s,g}\|_{H^s} \leq \frac{\delta}{2\sqrt{\lambda}}.$$

2 BACKGROUND THEOREMS

We shall use the following two general theorems.

THEOREM 4 *([5,2]) Let H_K be a Hilbert space admitting the reproducing kernel $K(p,q)$ on a set E. Let $L : H_K \to \mathcal{H}$ be a bounded linear operator on H_K into $\mathcal{H}$. For $\lambda > 0$ introduce the inner product in H_K and call it H_{K_λ} as*

$$\langle f_1, f_2 \rangle_{H_{K_\lambda}} = \lambda \langle f_1, f_2 \rangle_{H_K} + \langle Lf_1, Lf_2 \rangle_{\mathcal{H}}, \tag{2.8}$$

then H_{K_λ} is the Hilbert space with the reproducing kernel $K_\lambda(p,q)$ on E and satisfying the equation

$$K(\cdot, q) = (\lambda I + L^* L) K_\lambda(\cdot, q), \tag{2.9}$$

where L^* is the adjoint of $L : H_K \to \mathcal{H}$.

THEOREM 5 *([5,2]) Let H_K, L, $\mathcal{H}$, E and K_λ be as in Theorem 4. Then, for any $\lambda > 0$ and for any $g \in \mathcal{H}$, the extremal function in*

$$\inf_{f \in H_K} \left(\lambda \|f\|_{H_K}^2 + \|Lf - g\|_{\mathcal{H}}^2 \right) \tag{2.10}$$

exists uniquely and the extremal function is represented by

$$f_{\lambda,g}^*(p) = \langle g, LK_\lambda(.,p) \rangle_{\mathcal{H}} \tag{2.11}$$

which is the member of H_K attaining the infimum in (2.10).

3 PROOF OF THEOREM 1

First, of course, we have the inequality

$$\|HF\|_{L_2(\mathbf{R}^n)}^2 \leq \|F\|_{H^s}^2; \tag{3.12}$$

that is, the operator H is a bounded linear operator from H^s into $L_2(\mathbf{R}^n)$. Then we can see directly that

$$K_\lambda(x,y;H)$$

$$= \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} \frac{e^{ip\cdot(x-y)}}{\lambda(|p|^2+1)^s + (p_n^2 + |p'|^4)} dp \tag{3.13}$$

satisfies the functional equation (2.9) in our situation; that is, it is the reproducing kernel for the Hilbert space with the norm square

$$\lambda\|F\|_{H^s}^2 + \|HF\|_{L_2(\mathbf{R}^n)}^2.$$

In particular, we thus obtain (1.5) from Theorem 5.

In order to prove the result (1.6), as we see from (1.3) note that any member $F \in H^s$ is represented uniquely by a function $\mathbf{F}$ in the form

$$F(x) = \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} \frac{e^{ix\cdot\eta}}{(1+|\eta|^2)^s} \mathbf{F}(\eta) d\eta \tag{3.14}$$

satisfying

$$\frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} \frac{1}{(1+|\eta|^2)^s} |\mathbf{F}(\eta)|^2 d\eta < \infty$$

and

$$\|F\|_{H^s}^2 = \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} \frac{1}{(1+|\eta|^2)^s} |\mathbf{F}(\eta)|^2 d\eta. \tag{3.15}$$

Then, we insert this F in (1.1) and we have $(HF)(x)$. Then, we take it as $g(\xi)$ in (1.5) and we obtain, directly

$$F^*_{\lambda,s,g}(x)$$

$$= \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} \frac{(\eta_n^2 + |\eta'|^4) e^{ix\cdot\eta}}{(\lambda(1+|\eta|^2)^s + (\eta_n^2 + |\eta'|^4))(1+|\eta|^2)^s} \mathbf{F}(\eta) d\eta. \tag{3.16}$$

From (3.14) and (3.16) we thus obtain the desired result (1.6).

4 PROOF OF THEOREM 2

As in the proof of the weighted convolution inequalities in [4], we obtain directly, for $p = 2$ and $\rho_2 \equiv 1$

$$\int_{\mathbf{R}^n} |F^*_{\lambda,s,g}(x)|^2 dx$$

$$\leq \int_{\mathbf{R}^n} e^{-|\xi|^2} d\xi \int_{\mathbf{R}^n} \int_{\mathbf{R}^n} |g(\xi)|^2 e^{|\xi|^2} |Q_{\lambda,s}(\xi - x)|^2 d\xi dx$$

$$= \int_{\mathbf{R}^n} e^{-|\xi|^2} d\xi \int_{\mathbf{R}^n} |g(\xi)|^2 e^{|\xi|^2} d\xi \int_{\mathbf{R}^n} |Q_{\lambda,s}(\xi)|^2 d\xi$$

$$\le \pi^{n/2}\int_{\mathbf{R}^n}|g(\xi)|^2 e^{|\xi|^2}d\xi\cdot\frac{1}{4\lambda(2\pi)^n}\int_{\mathbf{R}^n}\frac{dp}{(|p|^2+1)^s}$$
$$=\frac{\Gamma(s-n/2)}{2^{n+2}\Gamma(s)}\frac{1}{\lambda}\int_{\mathbf{R}^n}|g(\xi)|^2 e^{|\xi|^2}d\xi.$$

5 PROOF OF THEOREM 3

We first have

$$\begin{aligned}\hat{F}^*_{\lambda,s,g_\delta}(p) &= \hat{g}_\delta(p)\frac{(-ip_n+|p'|^2)}{\lambda(|p|^2+1)^s+(p_n^2+|p'|^4)},\\ \hat{F}^*_{\lambda,s,g}(p) &= \hat{g}(p)\frac{(-ip_n+|p'|^2)}{\lambda(|p|^2+1)^s+(p_n^2+|p'|^4)}.\end{aligned}$$

It follows that

$$\hat{F}^*_{\lambda,s,g_\delta}(p)-\hat{F}^*_{\lambda,s,g}(p)=\frac{(\hat{g}_\delta(p)-\hat{g}(p))(-ip_n+|p'|^2)}{\lambda(|p|^2+1)^s+(p_n^2+|p'|^4)}.$$

Hence, we obtain

$$\begin{aligned}(1+|p|^2)^s|(\hat{F}^*_{\lambda,s,g_\delta}(p)-\hat{F}^*_{\lambda,s,g}(p))|^2 &\le \frac{|\hat{g}_\delta(p)-\hat{g}(p)|^2(1+|p|^2)^s}{4\lambda(|p|^2+1)^s}\\ &= \frac{|\hat{g}_\delta(p)-\hat{g}(p)|^2}{4\lambda}.\end{aligned}$$

Integrating the latter inequality over $\mathbf{R}^n$ we obtain the desired result.

6 LIMITING PROPERTIES

Our solution (1.5) will give a practical formula for the inhomogeneous heat equation. We will show experimental results by using computers. There, we will see that in order to overcome the difficulty in the equation, we must work hardly; that is, we must take a very small λ and we must calculate the integral (1.5) hardly in the sense of numerical. Computers help us this hard work to calculate the integral for a very small λ.

Meanwhile, for any $\lambda > 0$, we shall define a linear mapping

$$M_{\lambda,s}: L_2(\mathbf{R}^n)\to H^s$$

by $M_{\lambda,s}(g) = F^*_{\lambda,s,g}$. Now, we consider the composite operators $HM_{\lambda,s}$ and $M_{\lambda,s}H$. Using Fourier's integrals it can be shown that for $F \in H^s$,

$$\begin{aligned}(M_{\lambda,s}HF)(x) &= \frac{1}{(2\pi)^n}\int_{\mathbf{R}^n}\Big\{F(\xi) \\ &\cdot\int_{\mathbf{R}^n}\frac{(p_n^2+|p'|^4)e^{-ip\cdot(\xi-x)}dp}{\lambda(|p|^2+1)^s+(p_n^2+|p'|^4)}\Big\}d\xi \end{aligned} \tag{6.17}$$

and for $g \in L_2(\mathbf{R}^n)$,

$$\begin{aligned}(HM_{\lambda,s}g)(x) &= \frac{1}{(2\pi)^n}\int_{\mathbf{R}^n}\Big\{g(\xi) \\ &\cdot\int_{\mathbf{R}^n}\frac{(p_n^2+|p'|^4)e^{-ip\cdot(\xi-x)}dp}{\lambda(|p|^2+1)^s+(p_n^2+|p'|^4)}\Big\}d\xi. \end{aligned} \tag{6.18}$$

Setting

$$\Delta_{\lambda,s}(x-\xi) = \frac{1}{(2\pi)^n}\int_{\mathbf{R}^n}\frac{(p_n^2+|p'|^4)e^{-ip\cdot(\xi-x)}}{\lambda(|p|^2+1)^s+(p_n^2+|p'|^4)}dp$$

in (6.17) and (6.18), we have

$$(M_{\lambda,s}HF)(x) = \int_{\mathbf{R}^n}F(\xi)\Delta_{\lambda,s}(x-\xi)d\xi, \quad (F \in H^s) \tag{6.19}$$

and

$$(HM_{\lambda,s}g)(x) = \int_{\mathbf{R}^n}g(\xi)\Delta_{\lambda,s}(x-\xi)d\xi, \quad (g \in L_2(\mathbf{R}^n)). \tag{6.20}$$

Then we obtain that

$$\lim_{\lambda\to 0}\Delta_{\lambda,s}(x-\xi) = \delta(x-\xi), \tag{6.21}$$

$$\lim_{\lambda\to 0}M_{\lambda,s}H = I \tag{6.22}$$

and

$$\lim_{\lambda\to 0}HM_{\lambda,s} = I. \tag{6.23}$$

The precise mean of (6.19) and (6.22) is given as follows: For any $F \in H^s$

$$\lim_{\lambda \to 0} (M_{\lambda,s} HF)(x) = F(x) \tag{6.24}$$

uniformly on $\mathbf{R}^n$(cf. [7], Section 3). The precise mean of (6.20) and (6.23) is given as follows: For any $g \in \mathcal{R}(H) + \mathcal{R}(H)^{\perp}$

$$\lim_{\lambda \to 0} HM_{\lambda,s} g = g$$

in $L_2(\mathbf{R}^n)$(cf. [7]).

7 NUMERICAL EXPERIMENTS WITH FIGURES

Now we give experimental results to see the behaviour of

$$\lim_{\lambda \to 0} HM_{\lambda,s}$$

on $L_2(\mathbf{R}^2) \setminus \mathcal{R}(\Box_c)$. Here, we consider $g(x) = \chi_{[-1,1]}(x_1) \times \chi_{[-1,1]}(x_2)$ then $g \in L_2(\mathbf{R}^2) \setminus \mathcal{R}(\Box_c)$. See Figures 1 - 4.

Similarly, we consider the function $g(x) = e^{-|x|^2}(n = 2)$ then $g \in H^s$. We see from Figures 5 - 8 that

$$\lim_{\lambda \to 0} (HM_{\lambda,s} g)(x) = g(x).$$

In all the cases, we assume that

$$n = 2 \quad (x = (x_1, x_2) = (x, t)).$$

Further, space $x_1 = x$ is the right hand side direction and time $x_2 = t$ is the deep direction.

The above Figures are

$$F^*_{\lambda,2,g}(x_1, x_2)$$

and the below Figures are

$$HF^*_{\lambda,2,g}(x_1, x_2).$$

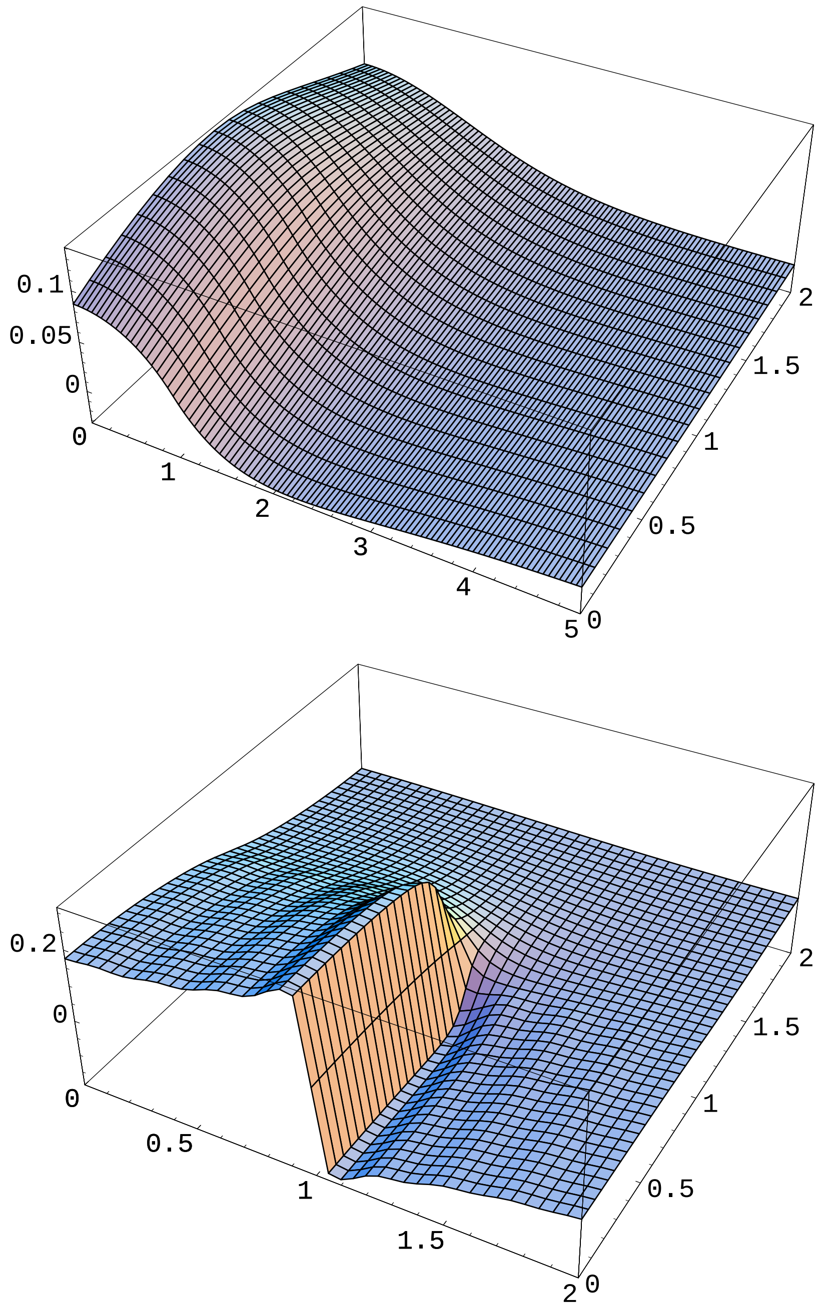


Figure 1: For $g(x_1, x_2) = \chi_{[-1,1]}(x_1) \times \chi_{[-1,1]}(x_2)$ on $\mathbf{R}^2$, the figures of $F^*_{\lambda,2,g}(x_1, x_2)$ and $HF^*_{\lambda,2,g}(x_1, x_2)$ for $\lambda = 10^0$.

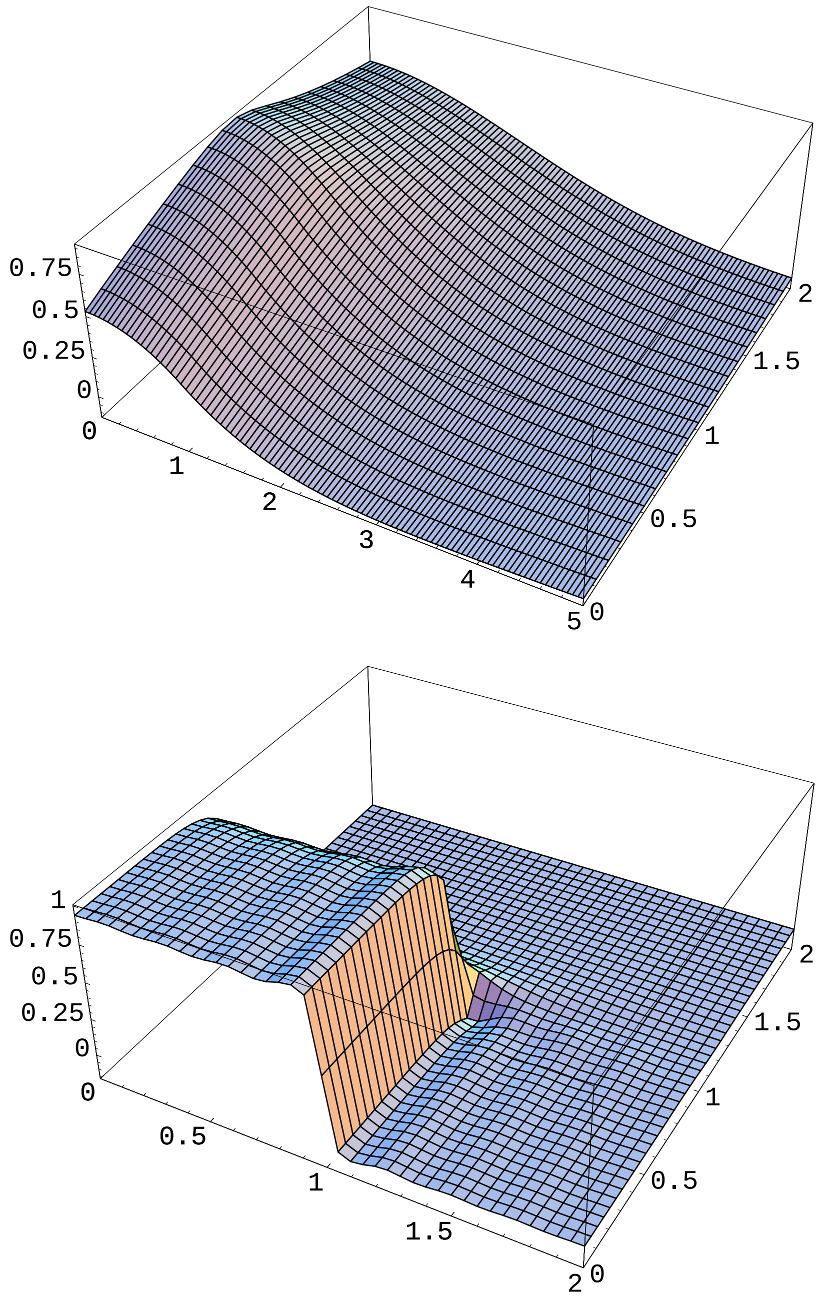


Figure 2: For $g(x_1, x_2) = \chi_{[-1,1]}(x_1) \times \chi_{[-1,1]}(x_2)$ on $\mathbf{R}^2$, the figures of $F^*_{\lambda,2,g}(x_1, x_2)$ and $HF^*_{\lambda,2,g}(x_1, x_2)$ for $\lambda = 10^{-2}$.

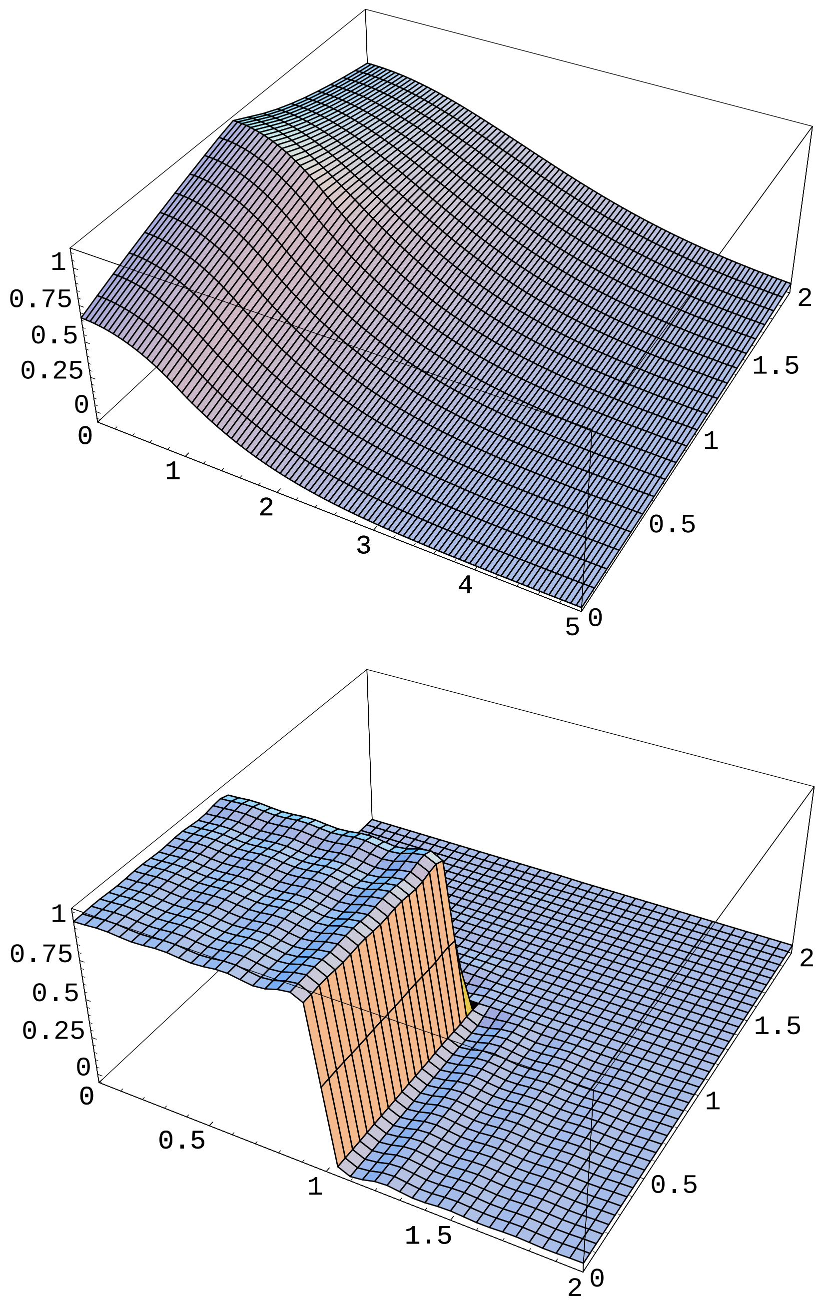


Figure 3: For $g(x_1, x_2) = \chi_{[-1,1]}(x_1) \times \chi_{[-1,1]}(x_2)$ on $\mathbf{R}^2$, the figures of $F^*_{\lambda,2,g}(x_1, x_2)$ and $HF^*_{\lambda,2,g}(x_1, x_2)$ for $\lambda = 10^{-4}$.

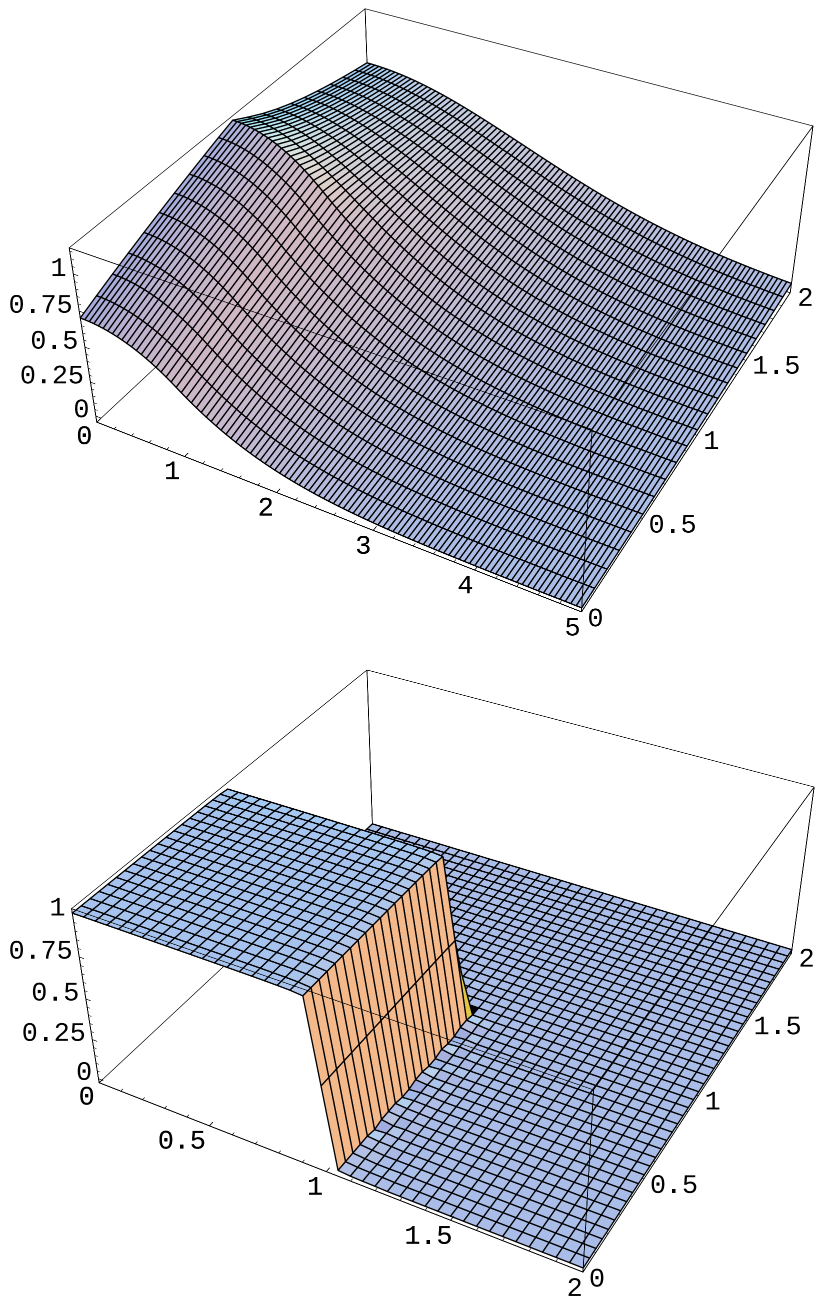
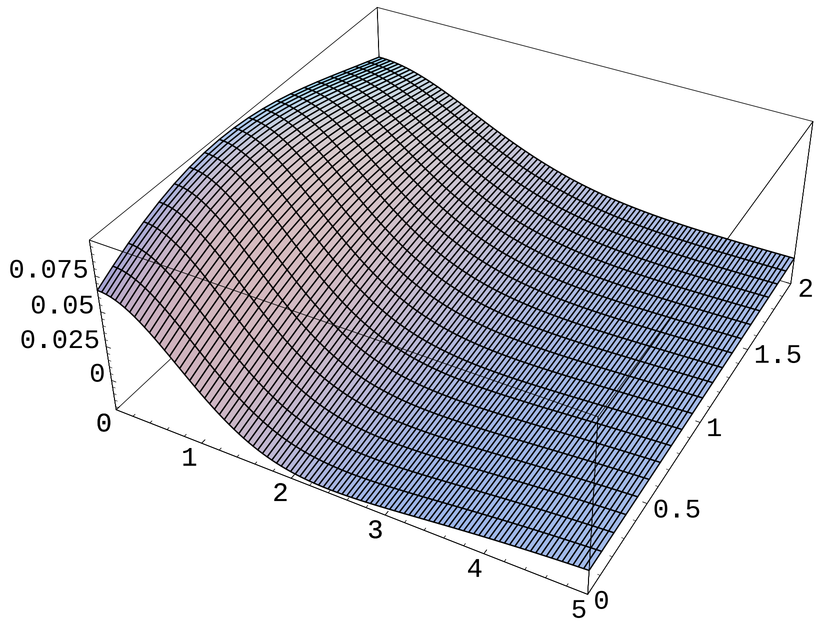


Figure 4: For $g(x_1, x_2) = \chi_{[-1,1]}(x_1) \times \chi_{[-1,1]}(x_2)$ on $\mathbf{R}^2$, the figures of $F^*_{\lambda,2,g}(x_1, x_2)$ and $HF^*_{\lambda,2,g}(x_1, x_2)$ for $\lambda = 10^{-6}$.



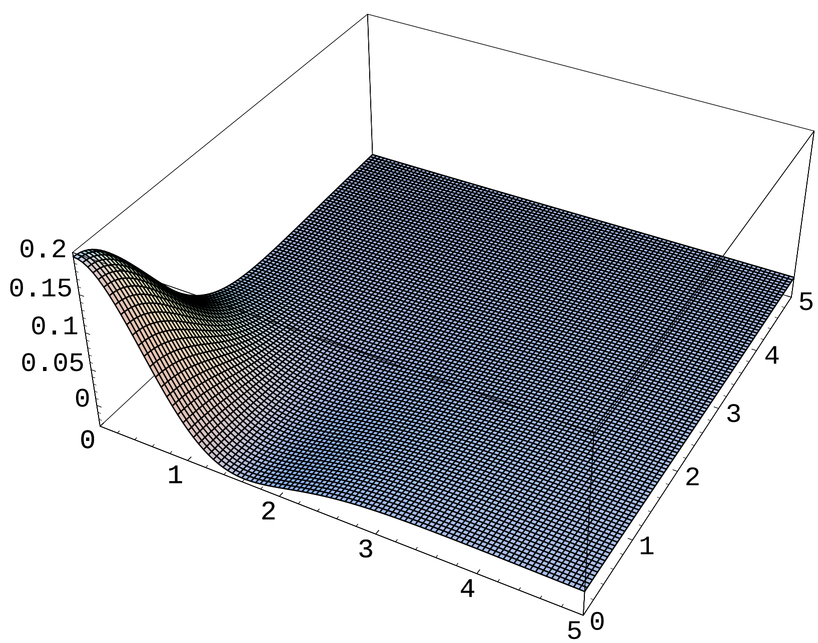


Figure 5: For $g(x_1, x_2) = e^{-(x_1^2+x_2^2)}$ on $\mathbf{R}^2$, the figures of $F^*_{\lambda,2,g}(x_1, x_2)$ and $HF^*_{\lambda,2,g}(x_1, x_2)$ for $\lambda = 10^0$.

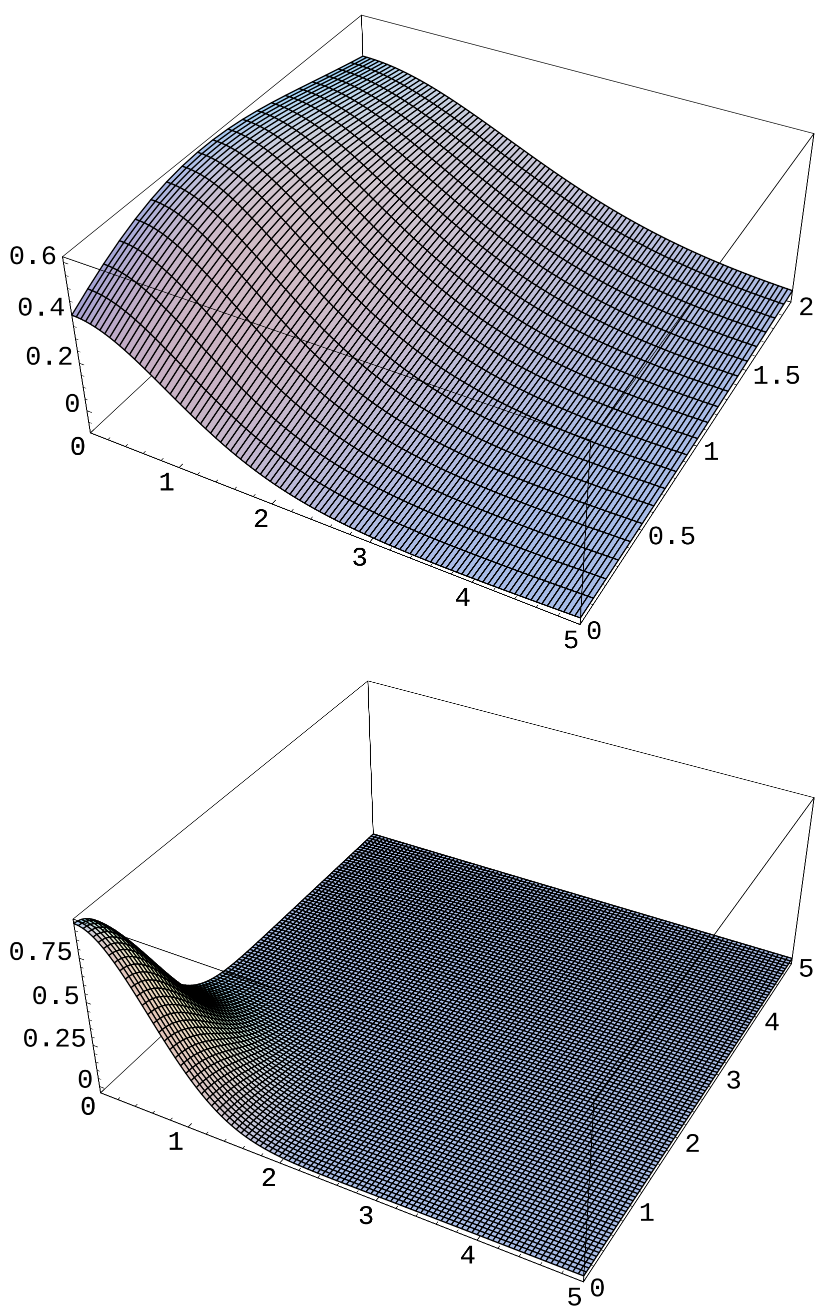


Figure 6: For $g(x_1, x_2) = e^{-(x_1^2+x_2^2)}$ on $\mathbf{R}^2$, the figures of $F^*_{\lambda,2,g}(x_1, x_2)$ and $HF^*_{\lambda,2,g}(x_1, x_2)$ for $\lambda = 10^{-2}$.

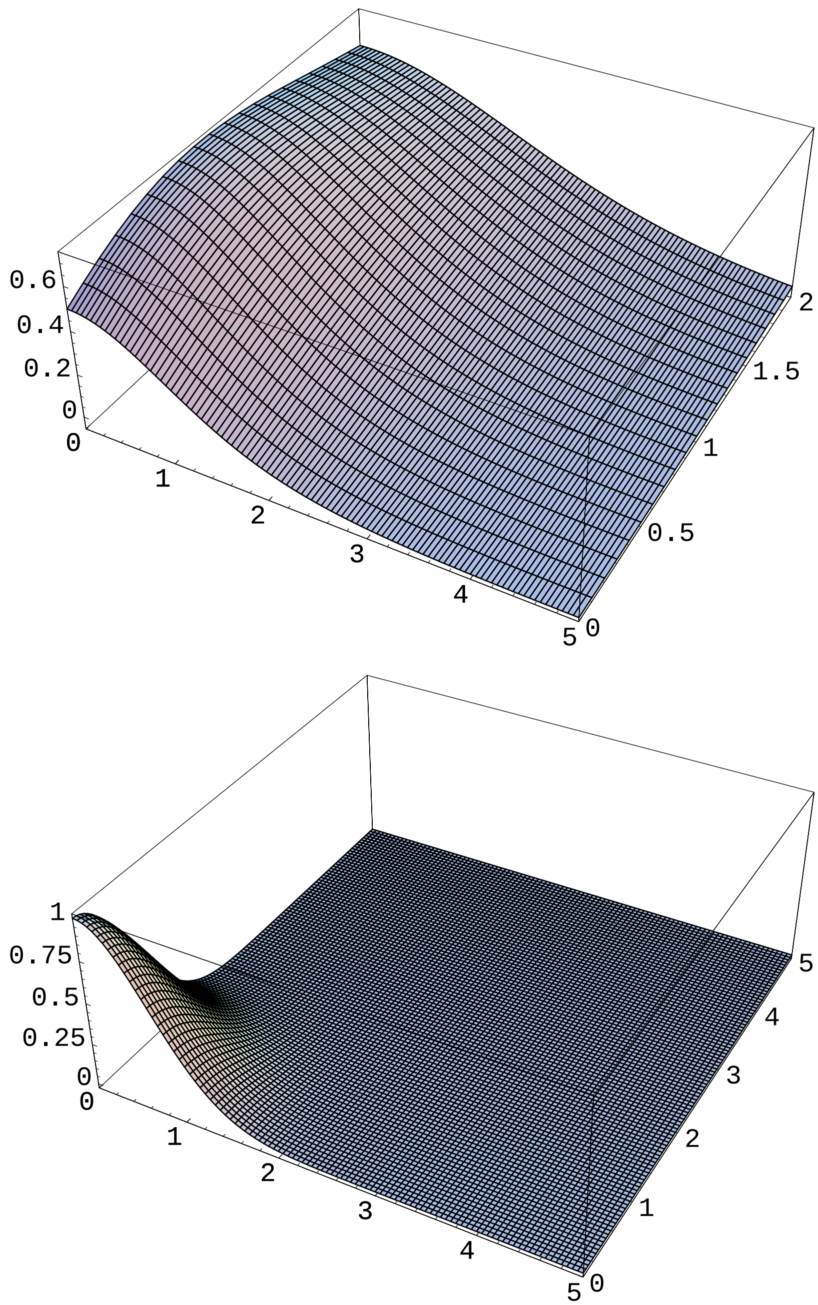


Figure 7: For $g(x_1, x_2) = e^{-(x_1^2+x_2^2)}$ on $\mathbf{R}^2$, the figures of $F^*_{\lambda,2,g}(x_1, x_2)$ and $HF^*_{\lambda,2,g}(x_1, x_2)$ for $\lambda = 10^{-4}$.

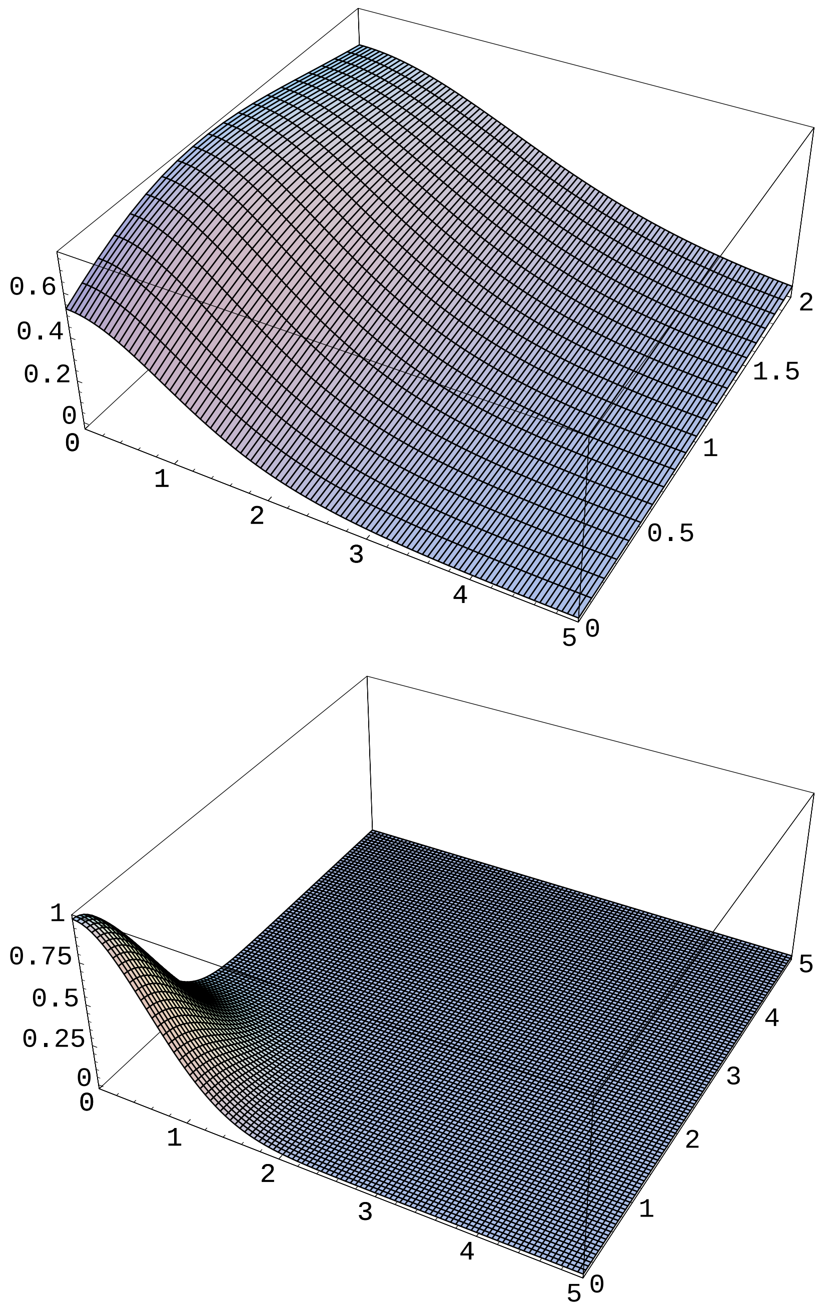


Figure 8: For $g(x_1, x_2) = e^{-(x_1^2+x_2^2)}$ on $\mathbf{R}^2$, the figures of $F^*_{\lambda,2,g}(x_1, x_2)$ and $HF^*_{\lambda,2,g}(x_1, x_2)$ for $\lambda = 10^{-6}$.

Acknowledgements

This research of the second author was partially supported by the Grant-in-Aid for the Scientific Research (C)(2) (No. 16540137) from the Japan Society for the Promotion Science and by the Mitsubishi Foundation, the 36th, Natural Sciences, No. 20 (2005-2006).

References

[1] M. Asaduzzaman, T. Matsuura, and S. Saitoh, *Constructions of approximate solutions for linear differential equations by reproducing kernels and inverse problems*, Advances in Analysis, Proceedings of the 4th International ISAAC Congress (2005), 335-344. (World Scientific).

[2] D-W, Byun and S. Saitoh, *Best approximation in reproducing kernel Hilbert spaces*, Proc. of the 2nd International Colloquium on Numerical Analysis, VSP-Holland, 55–61(1994).

[3] S. Saitoh, *Integral Transforms, Reproducing Kernels and their Applications*, Pitman Res. Notes in Math. Series **369**(1997), Addison Wesley Longman Ltd, UK.

[4] S. Saitoh, *Weighted L_p-norm inequalities in convolutions*, Survey on Classical Inequalities (T. M. Rassias, ed.), Kluwer Academic Publishers, 2000, pp. 225-234.

[5] S. Saitoh, *Approximate Real Inversion Formulas of the Gaussian Convolution*, Applicable Analysis, **83**(2004, 727-733.

[6] S. Saitoh, *Applications of Reproducing Kernels to Best Approximations, Tikhonov Regularizations and Inverse Problems*, Advances in Analysis, Proceedings of the 4th International ISAAC Congress (2005), (World Scientific), 439-445.

[7] S. Saitoh, *Best approximation, Tikhonov regularization and reproducing kernels*, Kodai. Math. J., **28**(2005), 359-367.

JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS,VOL.5,NO.3,255-262,2007,COPYRIGHT 2007 EUDOXUS PRESS,LLC

On the Periods of Some Transformation Used in Digital Image Scrambling and the Calculation Method to These Reverse Transformation

Shaochun Fu[1,3]; Jiansheng Cao[2]; Chengxian Xu[3]
[1]Xi'an Jiaotong University,Xi'an
[2]Department of Mathematics,East Tennessee Sate University, Johnson City, TN 37614
[3]School of Economics and Management Beijing Jiaotong University

Abstract It is known that image compression is very active research field. Aronad Transform, Permutation Transform and Fibonacci Transformation are common methods to be used in the image compression. But the computations of periods of these transforms are very complicated and inverse transforms of the original images are difficult to compute. In this paper, we analyzed several transformations used in digital image scrambling and obtained some results on the periods of these transformations to make the computing complexity lower than the other when the scrambling image deciphered. The fast calculation of the reverse transformations is also given in the paper.

Keywords linear transformation; affine transformation; integral residue rings, period; digital image scrambling

1. Introduction

With the rapid development of multimedia technology, more and more digital images are transferred on the network. Because this image information possibly involves with national security or company benefits, its value becomes more and more important for the researchers. So it is extremely important for the security of image transmission on the network and image encryption. During the process of digital image scrambling, generally the original image is pretreated is the first step , and the common method for pretreatment is to make proper transformation of original image. There are many methods for image transformation, such as Aronld transformation [1], Permutation transformation [2], Fibonacci transformation [3] and affine transformation [4]. The calculations of periods of the above transforms are very complicated. So far the computations of all periods of these transforms have not been solved completely. In this paper, we give the computational methods to calculate the periods of these transforms and solve the problems proposed in [4] to estimate the calculation of periods for different size images by using affine transform.

Because inverse transformation of every transform is not easy o calculate, the fast calculation of the inverse transform is given in this paper. The period of Aronld transform is 192, lots of unnecessary computation is added during decryption process of Aronld transform. So we need to find the method to calculate the periods of these transforms to lower the computational complexity. In this paper, we give a method that can complete the above transforms in one step. Therefore, these four transforms are very useful in the image compression. We assume that $N = 2^e$.

2. Polynomial and its property on the ring

First let's take a look at polynomial [5] on the ring of $Z/2^e$.

$$f(x) = x^n + \sum_{j=0}^{n-1} c_j x^j, \tag{1}$$

Where $c_j = \sum\limits_{i=0}^{e-1} C_{j,i} 2^i$, $f(x)$ also has binary dissociation:

$$f(x) = \sum_{i=0}^{e-1} f_i(x) 2^i, \tag{2}$$

Here $f_i = \sum\limits_{j=0}^{n} c_{j,i}(x) x^i$,

From [5], the maximum period of $f(x)$ is $2^{e-1}(2^n - 1)$.

$f(x)$ has $2^{e-1}(2^n-1)$ as its period if and only if (1) $f_0(x)$ if F_2's prime polynomial, (2)$\Delta_f \neq 0(modf_0(x), 2)$, here

$$\Delta_f = \begin{cases} h_f(x) & (modf_0(x), 2), \quad e = 2, \\ h_f(x) \quad (h_f(x)+1), & (modf_0(x), 2), \quad e \geq 3 \end{cases} \tag{3}$$

h$_f(x) = r(x) + q(x) f_1(x)(modf_0(x), 2)$ (4)

and $r(x)$, $q(x)$ satisfy

$$x^{2^n - 1} = f_0(x) q(x) + 2r(x) mod 2^e. \tag{5}$$

3. Determination of periods

(1) Aronld Transform

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} mod\ N, \qquad x, y \epsilon \{0, 1, 2, ..., N-1\} \tag{6}$$

Here,N is the pixel number of image's height and width.

Let T= $\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$,

Then we have

$$\left| \; T - Ix \; \right| = \begin{vmatrix} 1-x & 1 \\ 1 & 2-x \end{vmatrix} = x^2 - 3x + 1$$

Therefore

$$\left| \; T - Ix \; \right| mod2 = x^2 + x + 1,$$

From

$x^3 - 1 = (x-1)(x^2 + x + 1)$
Then q(x)=x-1,r(x)=0 Again in $x^2 + x + 1, f_1 = 0,$

so $h_{f(x)} = 0,\ namely,\ \Delta_f = 0,$

Therefore,the maximum period of $x^2 - 3x + 1$ is $2^6 \times 3 = 192$. T's change tendency is as following:

T	T^2	T^4	T^{64}	T^{128}	T^{192}	T^{384}
$\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$	$\begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix}$	$\begin{pmatrix} 13 & 21 \\ 21 & 34 \end{pmatrix}$	$\begin{pmatrix} 29 & 197 \\ 197 & 226 \end{pmatrix}$	$\begin{pmatrix} 226 & 59 \\ 59 & 29 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

(2)Permutation transformation

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} mod\ N, \qquad x, y\epsilon\{0, 1, 2, ..., N-1\} \qquad (7)$$

Where, $(ad - bc = 1, a, b, c, d\epsilon Z)$.When a=b=c=1,d=2, Permutation transformation is Aronld transformation. When a=b=c=1,d=0, Permutation transformation is Fibonacci transformation.

Let

$T_1 = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$,

Then

$$\left| \; T - Ix \; \right| = \begin{vmatrix} a-x & b \\ c & d-x \end{vmatrix} = x^2 - (a+d)x + (ad - bc) = x^2 - (a+d)x + 1 \quad (8)$$

So, when a+d is an odd number, the maximum period of T_1 on the ring F_2^8 is 384. When a+d is an even number, $|T_1 - Ix| mod2 = x^2 + 1$ then the period of T_1 on the ring F_2^8 is the multiplier of 256.
Example 1. In Permutation Transformation, when assumes

T=$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$, $\quad a = c = d = 1, b = 0,$

We have

$\mathrm{T}^{256} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$

Degradation is occurred for T's period.
When assumes

T= $\begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$

then $T^{128} = I$

When assumes

T= $\begin{pmatrix} 1 & 0 \\ 4 & 1 \end{pmatrix}$

then $T^{64} = I...;$

Assumes

T= $\begin{pmatrix} 1 & 0 \\ 128 & 0 \end{pmatrix}$

then $T^2 = I$

So all Ts in the above examples satisfy with conditions ad-bc=1.$(0, 0)$is their common stationary point.
Example 1 illustrates that, in the permutation transformation, the choices of a, b, c, d are related each other. So we suggest that when a, d is selected, it should be satisfied with a+d=1mod2

(3) Fibonacci transformation

In Fibonacci transformationtransformation matrix is:

T= $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$

then

$$\left| \; T - Ix \; \right| = \begin{vmatrix} 1-x & 1 \\ 1 & -x \end{vmatrix} = x^2 - x + 1$$

so $|T - Ix|$mod2=$x^2 + x + 1$, while $\Delta_f = 0$, so, the period$(when\ N = 2^e)$ of Fibonacci transformation is the multiplier of $3 \times 2^{e-2}$,when e=8, its period is smaller than 192.

(4) Affine transformation

The general form of affine transformation is:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix} \tag{9}$$

The affine transformation is used in text [4] is:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & N-1 \\ N-1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} modN, \quad x, y\epsilon\{0, 1, 2, ..., N-1\} \tag{10}$$

This affine transformation is combined by two parts, they are marked as T(1), T(2) respectively,
namely:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = T^{(1)}\begin{pmatrix} x \\ y \end{pmatrix} + T^{(2)}, Here \ \ T^{(1)} = \begin{pmatrix} 1 & N-1 \\ N-1 & 0 \end{pmatrix}, \ T^{(1)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Within F_2^8 , the period of T(2) is 256, while

$$\left| \ T^{(1)} - Ix \ \right| = \begin{vmatrix} 1-x & N-1 \\ N-1 & -x \end{vmatrix} = x^2 - x - (N-1)^2 = x^2 - x - 255^2$$

Thus

$$\left| \ T^{(1)} - Ix \ \right| mod2 = x^2 + x + 1$$

Since $\Delta_f \neq 0$, the period of $T^{(1)}$ is (withinF_2^8)384. From the following conclusion, for the affine transformation given by formulation (10), its period is 384, thus the problem of the period of affine transformation is smaller than the period of Aronld transformation doesn't exist. Of course, we also solve the period estimation problem of affine transformation to all kinds of image size that proposed by [4].

In affine transformation

$$P\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix} modN \tag{11}$$

assumes that the period of $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is N_0, B=A^{N_0-1}+A^{N_0-2}+...+A+I

then

$$AB = A^{N_0} + A^{N_0-1} + ... + A^2 + A = I + A^{N_0-1} + A^{N_0-2} + ... + A^2 + A = B \tag{12}$$

$$P^2\begin{pmatrix} x \\ y \end{pmatrix} = A\left(A\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix} \right) + \begin{pmatrix} e \\ f \end{pmatrix} = A^2\begin{pmatrix} x \\ y \end{pmatrix} + A\begin{pmatrix} e \\ f \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix} \tag{13}$$

$$\begin{aligned} \mathrm{P}^T\begin{pmatrix} x \\ y \end{pmatrix} &= A^T\begin{pmatrix} x \\ y \end{pmatrix} + A^{T-1}\begin{pmatrix} e \\ f \end{pmatrix} + A^{T-2}\begin{pmatrix} e \\ f \end{pmatrix} + ... + A\begin{pmatrix} e \\ f \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix} \\ &= A^T\begin{pmatrix} x \\ y \end{pmatrix} + [A^{T-1} + A^{T-2} + ... + A + I]\begin{pmatrix} e \\ f \end{pmatrix} \end{aligned} \tag{14}$$

$$P^{N_0}\begin{pmatrix} x \\ y \end{pmatrix} = A^{N_0}\begin{pmatrix} x \\ y \end{pmatrix} + B\begin{pmatrix} e \\ f \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} + B\begin{pmatrix} e \\ f \end{pmatrix} \tag{15}$$

$$P^{N_0+1}\begin{pmatrix} x \\ y \end{pmatrix} = A\begin{pmatrix} x \\ y \end{pmatrix} + B\begin{pmatrix} e \\ f \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix} \tag{16}$$

$$P^{N_0+2}\begin{pmatrix} x \\ y \end{pmatrix} = A^2\begin{pmatrix} x \\ y \end{pmatrix} + B\begin{pmatrix} e \\ f \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix} \tag{17}$$

$$\begin{aligned} \mathrm{P}^{2N_0}\begin{pmatrix} x \\ y \end{pmatrix} &= A^{N_0}\begin{pmatrix} x \\ y \end{pmatrix} + B\begin{pmatrix} e \\ f \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix} = A^{N_0}\begin{pmatrix} x \\ y \end{pmatrix} + 2B\begin{pmatrix} e \\ f \end{pmatrix} \\ &= \begin{pmatrix} x \\ y \end{pmatrix} + 2B\begin{pmatrix} e \\ f \end{pmatrix} \end{aligned} \tag{18}$$

in a similar way,for V k$\epsilon Z, there\ is$

$$P^{kN_0}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} + kB\begin{pmatrix} e \\ f \end{pmatrix} \text{(19)}$$

let

$$B\begin{pmatrix} e \\ f \end{pmatrix} = \begin{pmatrix} e' \\ f' \end{pmatrix} \tag{20}$$

Keeps min$\{gcd(e', N), gcd(f', N)\} = k'$,then when $k^0 = N/k'$,

k$_0 B\begin{pmatrix} e \\ f \end{pmatrix} = \begin{pmatrix} k_0 e' \\ k_0 f' \end{pmatrix} mod\ N = 0,$

Thus N_0k_0 is one of the P's periods.
Let's prove N_0k_0 is P's minimum period.
Assumes the minimum period of P is $T = N_1K_1$, then

$\begin{pmatrix} x \\ y \end{pmatrix} = p^T\begin{pmatrix} x \\ y \end{pmatrix} = A^T\begin{pmatrix} x \\ y \end{pmatrix} + (A^{T-1} + A^{T-2} + \Lambda + A + I)\begin{pmatrix} e \\ d \end{pmatrix}, \forall x, y \in \{1, 2, \Lambda, N\}$

Since x, y are random numbers, it should be
A$^T = I\ mod\ N, moreover$

$$(A^{T-1} + A^{T-2} + \Lambda + A + I)\begin{pmatrix} e \\ f \end{pmatrix} = 0\ mod\ N \tag{21}$$

Keeps $T = N_0K_1$, and then there is

(A$^{T-1} + A^{T-2} + \Lambda + A + I)\begin{pmatrix} e \\ f \end{pmatrix} = k_1(A^{N_0-1} + \Lambda + A + I)\begin{pmatrix} e \\ d \end{pmatrix} = k_1\begin{pmatrix} e_1' \\ f_1' \end{pmatrix} = 0$

so,$k_1 \geq k_0$,hence $T \geq N_0K_0$. Namely N_0K_0 is P's minimum period. Especially if there is inverse for A-I, B=0modN, now,P's period equals to A's period. When A -I is inverse,B=0modN,
From

$(\mathrm{A}^{T-1}+A^{T-2}+\Lambda+A+I)(A-I) = (I+A^{T-1}+\Lambda+A^2+A)-(A^{T-1}+A^{T-2}+\Lambda+A+I) = 0$

we have,when(A-I)mod N is reversible, there must be

$\sum_{i=0}^{T-1} A^i = 0$

In this case, P's period is A's period.
If (A-I) is not reversible, then, for

$\mathrm{B}\begin{pmatrix} e \\ F \end{pmatrix} = \begin{pmatrix} e' \\ f' \end{pmatrix}$

e', f' must be even numbers.
P' period is the multiplier of $3 \times 2^{2e-2}$. When e=8, P's period is the multiplier of 3×2^{14}. for the affine transformation given by formulation (10), when $N = 2^8$, since $N_0 = 384, T^{(1)} - I$ is reversible, hence the period of T defined by formulation (10) is 384.

4.Rapid calculation method of transformation

For affine transformation (11) , there is following relational expression

$\mathrm{P}^{T1+T2}\begin{pmatrix} X \\ Y \end{pmatrix} = A^{T1+T2}\begin{pmatrix} X \\ Y \end{pmatrix} + A^{T1}\begin{pmatrix} e_2 \\ f_2 \end{pmatrix} + \begin{pmatrix} e_1 \\ f_1 \end{pmatrix},$

of which,

$\begin{pmatrix} e_2 \\ f_2 \end{pmatrix}, \begin{pmatrix} e_1 \\ f_1 \end{pmatrix}$

satisfies

$\mathrm{P}^{T1}\begin{pmatrix} x \\ y \end{pmatrix} = A^{T1}\begin{pmatrix} x \\ y \end{pmatrix} + A^{T1}\begin{pmatrix} e_1 \\ f_1 \end{pmatrix}, \mathrm{P}^{T2}\begin{pmatrix} x \\ y \end{pmatrix} = A^{T2}\begin{pmatrix} x \\ y \end{pmatrix} + A^{T1}\begin{pmatrix} e_2 \\ f_2 \end{pmatrix}.$ For Aronld transformation, arrangement transformation and Fibonacci transformation, when one of the A's periods is known, reverse calculation of A can be done within one step. Assumes that one of the transformation periods is N_0, A-l can be obtained by pre-calculating A^{N0-1} . For the calculation of A^{N0-1}, the method is as follows:

Assumes $\mathrm{N}_0 - 1 = \sum_{I=0}^{t} 2^i a_i, a_i = 0, 1$

Let

$\mathrm{T}_0 = A, \Lambda, T_i = T_{i-1}^2, \Lambda, T_t = T_{t-1}^2.$

Then

$A^{N_0-1} = \prod_{i:ai=1} T_i$ For instance: during image treatment, when the pixel number of the image height and width is 256, $t \leq 7$. So the calculation of A^{N0-1} is fairly easy. During decryption, A^{N0-1} is directly employed in cryptograph image, it is not necessary to continuously function $N_0 - 1$ times by using A. in a similar way, when affine transformation (11) is used, P^{N0-1} can be pre-calculated, at this time N_0 is P's period.

5.Conclusion

In this paper, we studied the period problem of Aronld transformation, Permutation transformation, Fibonacci transformation and affine transformation as well the fast calculation method of transformation, especially the reverse calculation for various transformations can be completed by only one step, these results not only has the practical usage in the image restore, but also effectively reduces the computational complexity during image restore.

REFERENCES

1. Ll Wei-Qing ,Tan Jian-Rong ,Peng Qun-Shen. A study on the global recognition method based on patch structure. [J].Chinese J computer. 1998;21(8):753-758.

2. Wu Min-Sheng, Wang Jie-Sheng ,Liu Shen-Quan. Permuttation transform of images. [J].Chinese J computer. 1998216514-519.

3. Qi Dong-Xu. Zou Jianchen, Han Xiaoyou.A new class of scrambling transformation and it's application in the image information covering [J]Science in china(Series E),2000;43(3):304-312.

4. Bai Senl Cao Chang-xiul ,Bai Lin. A New Technology for Scrambling Digital Image Based on Affined Transformation. [J]Computer engineering and application.2002281074-78.

5. Zong-Duo Dai.Binary Sequences Derived from ML-Sequences over Rings I:Periods Minimal Polynomials. J.Cryptology(1992)5:193-207.

Instructions to Contributors
Journal of Concrete and Applicable Mathematics

A quartely international publication of Eudoxus Press, LLC, of TN.

Editor in Chief: George Anastassiou
Department of Mathematical Sciences
University of Memphis
Memphis, TN 38152-3240, U.S.A.

1. Manuscripts hard copies in triplicate, and in English, should be submitted to the Editor-in-Chief:

Prof.George A. Anastassiou
Department of Mathematical Sciences
The University of Memphis
Memphis,TN 38152, USA.
Tel.001. 901.678.3144
e-mail:ganastss@memphis.edu.

Authors may want to recommend an associate editor the most related to the submission to possibly handle it.
Also authors may want to submit a list of six possible referees, to be used in case we cannot find related referees by ourselves.
2. Manuscripts should be typed using any of TEX,LaTEX,AMS-TEX,or AMS-LaTEX and according to EUDOXUS PRESS, LLC. LATEX STYLE FILE. (VISIT www.msci.memphis.edu/~ganastss/jcaam / to save a copy of the style file.)They should be carefully prepared in all respects. Submitted copies should be brightly printed (not dot-matrix), double spaced, in ten point type size, on one side high quality paper 8(1/2)x11 inch. Manuscripts should have generous margins on all sides and should not exceed 24 pages.
3. Submission is a representation that the manuscript has not been published previously in this or any other similar form and is not currently under consideration for publication elsewhere. A statement transferring from the authors(or their employers,if they hold the copyright) to Eudoxus Press, LLC, will be required before the manuscript can be accepted for publication.The Editor-in-Chief will supply the necessary forms for this transfer.Such a written transfer of copyright,which previously was assumed to be implicit in the act of submitting a manuscript,is necessary under the U.S.Copyright Law in order for the publisher to carry through the dissemination of research results and reviews as widely and effective as possible.
4. The paper starts with the title of the article, author's name(s) (no titles or degrees), author's affiliation(s) and e-mail addresses. The affiliation should comprise the department, institution (usually university or company), city, state (and/or nation) and mail code.
The following items, 5 and 6, should be on page no. 1 of the paper.

5. An abstract is to be provided, preferably no longer than 150 words.
6. A list of 5 key words is to be provided directly below the abstract. Key words should express the precise content of the manuscript, as they are used for indexing purposes. The main body of the paper should begin on page no. 1, if possible.
7. All sections should be numbered with Arabic numerals (such as: 1. INTRODUCTION) .
Subsections should be identified with section and subsection numbers (such as 6.1. Second-Value Subheading).
If applicable, an independent single-number system (one for each category) should be used to label all theorems, lemmas, propositions, corrolaries, definitions, remarks, examples, etc. The label (such as Lemma 7) should be typed with paragraph indentation, followed by a period and the lemma itself.
8. Mathematical notation must be typeset. Equations should be numbered consecutively with Arabic numerals in parentheses placed flush right,and should be thusly referred to in the text [such as Eqs.(2) and (5)]. The running title must be placed at the top of even numbered pages and the first author's name, et al., must be placed at the top of the odd numbed pages.
9. Illustrations (photographs, drawings, diagrams, and charts) are to be numbered in one consecutive series of Arabic numerals. The captions for illustrations should be typed double space. All illustrations, charts, tables, etc., must be embedded in the body of the manuscript in proper, final, print position. In particular, manuscript, source, and PDF file version must be at camera ready stage for publication or they cannot be considered.
Tables are to be numbered (with Roman numerals) and referred to by number in the text. Center the title above the table, and type explanatory footnotes (indicated by superscript lowercase letters) below the table.
10. List references alphabetically at the end of the paper and number them consecutively. Each must be cited in the text by the appropriate Arabic numeral in square brackets on the baseline.
References should include (in the following order):
initials of first and middle name, last name of author(s)
title of article,
name of publication, volume number, inclusive pages, and year of publication.
Authors should follow these examples:

<u>Journal Article</u>

1. H.H.Gonska,Degree of simultaneous approximation of bivariate functions by Gordon operators, (journal name in italics) *J. Approx. Theory*, 62,170-191(1990).

<u>Book</u>

2. G.G.Lorentz, (title of book in italics) *Bernstein Polynomials* (2nd ed.), Chelsea,New York,1986.

<u>Contribution to a Book</u>

3. M.K.Khan, Approximation properties of beta operators,in(title of book in italics) *Progress in Approximation Theory* (P.Nevai and A.Pinkus,eds.), Academic Press, New York,1991,pp.483-495.

11. All acknowledgements (including those for a grant and financial support) should occur in one paragraph that directly precedes the References section.
12. Footnotes should be avoided. When their use is absolutely necessary, footnotes should be numbered consecutively using Arabic numerals and should be typed at the bottom of the page to which they refer. Place a line above the footnote, so that it is set

off from the text. Use the appropriate superscript numeral for citation in the text.
13. After each revision is made please again submit three hard copies of the revised manuscript, including in the final one. And after a manuscript has been accepted for publication and with all revisions incorporated, manuscripts, including the TEX/LaTex source file and the PDF file, are to be submitted to the Editor's Office on a personal-computer disk, 3.5 inch size. Label the disk with clearly written identifying information and properly ship, such as:

Your name, title of article, kind of computer used, kind of software and version number, disk format and files names of article, as well as abbreviated journal name.

Package the disk in a disk mailer or protective cardboard. Make sure contents of disks are identical with the ones of final hard copies submitted!

Note: The Editor's Office cannot accept the disk without the accompanying matching hard copies of manuscript. No e-mail final submissions are allowed! The disk submission must be used.

14. Effective 1 Nov. 2005 the journal's page charges are $10.00 per PDF file page. Upon acceptance of the paper an invoice will be sent to the contact author. The fee payment will be due one month from the invoice date. The article will proceed to publication only after the fee is paid. The charges are to be sent, by money order or certified check, in US dollars, payable to Eudoxus Press, LLC, to the address shown in the Scope and
Prices section.

No galleys will be sent and the contact author will receive one(1) electronic copy of the journal issue in which the article appears.

15. This journal will consider for publication only papers that contain proofs for their listed results.

TABLE OF CONTENTS,JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS,VOL.5,NO.3,2007

VOLUME 5,NUMBER 4 OCTOBER 2007

ISSN:1548-5390 PRINT,1559-176X ONLINE

JOURNAL
OF CONCRETE
AND APPLICABLE
MATHEMATICS

EUDOXUS PRESS,LLC

SCOPE AND PRICES OF THE JOURNAL
Journal of Concrete and Applicable Mathematics

A quartely international publication of **Eudoxus Press,LLC**

Editor in Chief: George Anastassiou
Department of Mathematical Sciences,
University of Memphis
Memphis, TN 38152, U.S.A.
ganastss@memphis.edu

The main purpose of the "Journal of Concrete and Applicable Mathematics" is to publish high quality original research articles from all subareas of Non-Pure and/or Applicable Mathematics and its many real life applications, as well connections to other areas of Mathematical Sciences, as long as they are presented in a Concrete way. It welcomes also related research survey articles and book reviews.A sample list of connected mathematical areas with this publication includes and is not restricted to: Applied Analysis, Applied Functional Analysis, Probability theory, Stochastic Processes, Approximation Theory, O.D.E, P.D.E, Wavelet, Neural Networks,Difference Equations, Summability, Fractals, Special Functions, Splines, Asymptotic Analysis, Fractional Analysis, Inequalities, Moment Theory, Numerical Functional Analysis,Tomography, Asymptotic Expansions, Fourier Analysis, Applied Harmonic Analysis, Integral Equations, Signal Analysis, Numerical Analysis, Optimization, Operations Research, Linear Programming, Fuzzyness, Mathematical Finance, Stochastic Analysis, Game Theory, Math.Physics aspects, Applied Real and Complex Analysis, Computational Number Theory, Graph Theory, Combinatorics,
Computer Science Math.related topics,combinations of the above, etc. In general any kind of Concretely presented Mathematics which is Applicable fits to the scope of this journal.
Working Concretely and in Applicable Mathematics has become a main trend in many recent years,so we can understand better and deeper and solve the important problems of our real and scientific world.
"Journal of Concrete and Applicable Mathematics" is a peer- reviewed International Quarterly Journal.
We are calling for papers for possible publication. The contributor should send three copies of the contribution to the editor in-Chief typed in TEX, LATEX double spaced. [See: Instructions to Contributors]

Journal of Concrete and Applicable Mathematics(JCAAM)
ISSN:1548-5390 PRINT, 1559-176X ONLINE.
is published in January,April,July and October of each year by
EUDOXUS PRESS,LLC,
1424 Beaver Trail Drive,Cordova,TN38016,USA,
Tel.001-901-751-3553
anastassioug@yahoo.com
http://www.EudoxusPress.com.
Visit also www.msci.memphis.edu/~ganastss/jcaam.
Webmaster:Ray Clapsadle

Annual Subscription Current Prices:For USA and Canada,Institutional:Print $250,Electronic $220,Print and Electronic $310.Individual:Print $77,Electronic $60,Print

&Electronic $110.For any other part of the world add $25 more to the above prices for Print.
Single article PDF file for individual $8.Single issue in PDF form for individual $25.
The journal carries page charges $10 per page of the pdf file of an article,payable upon acceptance of the article within one month and before publication.
No credit card payments.Only certified check,money order or international check in US dollars are acceptable.
Combination orders of any two from JoCAAA,JCAAM,JAFA receive 25% discount,all three receive 30% discount.

JCAAM is reviewed and abstracted by AMS Mathematical Reviews,MATHSCI,and Zentralblaat MATH.

JCAAM IS A JOURNAL OF RAPID PUBLICATION

Editorial Board
Associate Editors

Institute of Mathematics
Gerhard Mercator University
D-47048 Duisburg
Germany
tel.0049-203-379-2676
knoop@math.uni-duisburg.de
Approximation Theory,Interpolation

16) Jerry Koliha
Dept. of Mathematics & Statistics
University of Melbourne
VIC 3010,Melbourne
Australia
koliha@unimelb.edu.au
Inequalities,Operator Theory,
Matrix Analysis,Generalized Inverses

17) Mustafa Kulenovic
Department of Mathematics
University of Rhode Island
Kingston,RI 02881,USA
kulenm@math.uri.edu
Differential and Difference Equations

18) Gerassimos Ladas
Department of Mathematics
University of Rhode Island
Kingston,RI 02881,USA
gladas@math.uri.edu
Differential and Difference Equations

19) Rupert Lasser
Institut fur Biomathematik & Biomertie,GSF
-National Research Center for environment
and health
Ingolstaedter landstr.1
D-85764 Neuherberg,Germany
lasser@gsf.de
Orthogonal Polynomials,Fourier Analysis,
Mathematical Biology

University of Rostock
D-18051 Rostock
Germany
manfred.tasche@mathematik.uni-rostock.de
Approximation Th.,Wavelet,Fourier Analysis,
Numerical Methods,Signal Processing,
Image Processing,Harmonic Analysis

35) Chris P.Tsokos
Department of Mathematics
University of South Florida
4202 E.Fowler Ave.,PHY 114
Tampa,FL 33620-5700,USA
profcpt@math.usf.edu,profcpt@chuma1.cas.usf.edu
Stochastic Systems,Biomathematics,
Environmental Systems,Reliability Th.

36) Lutz Volkmann
Lehrstuhl II fuer Mathematik
RWTH-Aachen
Templergraben 55
D-52062 Aachen
Germany
volkm@math2.rwth-aachen.de
Complex Analysis,Combinatorics,Graph Theory

JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS,VOL.5,NO.4,275-298,2007,COPYRIGHT 2007 EUDOXUS PRESS,LLC

Ito-Levy Processes

Tsvetelin Stefanov Zaevski
Department of Mathematics and Informatics
Sofia University "St. Kliment Ohridski"
Sofia, Bulgaria
e-mail: t_s_zaevski@abv.bg

November 15, 2006

Abstract

In this article we define the Ito-Levy processes, which are combinations of the Ito and Levy processes. The main idea is to consider the initial condition in a stochastic differential equation as a Levy process. Note the difference between the subordination and the Ito-Levy processes. While the substitution in the subordination is in the time variable, the substitution in the Ito-Levy processes is in the initial condition of the Ito process.

After that we shall examine the so defined processes and prove the formulas for the expectation, conditional expectation and the infinitesimal generator.

We shall also compare the propositions for the martingality when the process is of the Ito or of the Levy type and we shall show where the difference is when the process is of the Ito-Levy type.

In the next article we shall study the problem of changing the probability measure and construct an algorithm for constructing martingale measures.
Keywords: Ito-Levy process, expectations, infinitesimal generator, markov property, martingality.

1 Definition

Let X_t^x be an Ito process, which is a solution of the stochastic differential equation:

$$\begin{aligned} dX_t^x &= \mu\left(X_t^x\right)dt + \sigma\left(X_t^x\right)dB_t \\ X_0^x &= x, x \in \mathbf{R}, \end{aligned} \tag{1}$$

where B_t is a standard Brownian motion. The following theorem gives us sufficient conditions for the existence and uniqueness of the solution. (for proof see [6, theorem 5.2.1.])

Theorem 1.1. *Let $T > 0$, μ and σ be functions, which satisfy*

$$\begin{array}{l} |\mu(x)| + |\sigma(x)| \leq C(1 + |x|), \ \forall x \in \mathbf{R} \\ |\mu(x) - \mu(y)| + |\sigma(x) - \sigma(y)| \leq D|x - y|, \ \forall x, y \in \mathbf{R} \end{array}$$

for suitable constants C and D. Then the stochastic differential equation (1) *has a unique solution.*

Also let Y_t be a Levi process, independent of X_t, with a generating triple $(b, 0, \Pi)$. We denote the probability spaces of the processes X_t and Y_t by $\left(\Omega^X, \mathcal{F}^X, P^X\right)$ and $\left(\Omega^Y, \mathcal{F}^Y, P^Y\right)$ and the corresponding expectations by E^X and E^Y. Since the Levy processes are heavy tail processes, then when we examine the expectation $E(g(Y_t))$ we assume that $g(\cdot)$ is a suitable function.

We define

$$F_{t_1,t_2,...,t_n}(\mathrm{B_1}, \mathrm{B_2},..., \mathrm{B_n}) = \mathrm{E^Y}\left[\begin{array}{l} \mathrm{P^X}\left(X_{t_1}^{x_1} \in B_1, X_{t_2}^{x_2} \in B_2, ..., X_{t_n}^{x_n} \in B_n\right)\big| \\ \big|x_1 = Y_{t_1}^z, x_2 = Y_{t_2}^z, ..., x_n = Y_{t_n}^z \end{array}\right] \tag{2}$$

for arbitrary Borel sets $\mathrm{B_1}, \mathrm{B_2},..., \mathrm{B_n}$ We shall use the following Kolmogorov's theorem to prove that there exists a process Z_t^z with finite-dimensional distributions (2). (see [4])

Theorem 1.2 (Kolmogorov's extension theorem)**.** *For all $t_1, ..., t_n \in T$, $n \in \mathbf{N}$ let $\nu_{t_1,t_2,...,t_n}$ be a probability measures on R^n s.t.*

$$\nu_{t_{\alpha(1)},t_{\alpha(2)},...,t_{\alpha(n)}}(B_1 \times ... \times B_n) = \nu_{t_1,t_2,...,t_n}\left(B_{\alpha^{-1}(1)} \times ... \times B_{\alpha^{-1}(n)}\right) \tag{3}$$

for all permutations α on $\{1, 2, ..., n\}$ and

$$\nu_{t_1,t_2,...,t_n}(\mathrm{B_1} \times ... \times \mathrm{B_n}) = \nu_{t_1,t_2,...,t_n,t_{n+1},...,t_{n+m}}(\mathrm{B_1} \times ... \times \mathrm{B_n} \times R \times ... \times R) \tag{4}$$

for all $m \in \mathbf{N}$, where (of course) the set on the right hand side has a total of $n + m$ factors. Then there exists a probability space $(\Omega, \mathcal{F}, P)$ and a stochastic process $\{Z_t\}$ on Ω, $Z_t : \Omega \to \mathbf{R}$, s. t.

$$\nu_{t_1,t_2,...,t_n}(B_1 \times ... \times B_n) = P(Z_{t_1} \in B_1, ..., Z_{t_n} \in B_n)$$

for all $t_1, ..., t_n \in \mathbf{T}, n \in \mathbf{N}$ and all Borel sets B_i.

1. Verification of the agreement condition (3)

$$\begin{aligned} & F_{t_{\alpha(1)}, t_{\alpha(2)}, ..., t_{\alpha(n)}}(\mathrm{B_1}, \mathrm{B_2},..., \mathrm{B_n}) = \\ & \quad = \mathrm{E^Y}\left[\begin{array}{l} \mathrm{P^X}(X_{t_{\alpha(1)}}^{x_1} \in B_1, X_{t_{\alpha(2)}}^{x_2} \in B_2, ..., X_{t_{\alpha(n)}}^{x_n} \in B_n)\Big| \\ x_1 = Y_{t_{\alpha(1)}}^z, x_2 = Y_{t_{\alpha(2)}}^z, ..., x_n = Y_{t_{\alpha(n)}}^z \end{array}\right] = \\ & \quad = \mathrm{E^Y}\left[\begin{array}{l} \mathrm{P^X}\left(\begin{array}{l} X_{t_1}^{x_1} \in B_{\alpha^{-1}(1)}, X_{t_2}^{x_2} \in B_{\alpha^{-1}(2)}, ..., \\ X_{t_n}^{x_n} \in B_{\alpha^{-1}(n)} \end{array}\right)\Bigg| \\ x_1 = Y_{t_1}^z, x_2 = Y_{t_2}^z, ..., x_n = Y_{t_n}^z \end{array}\right] = \\ & \quad = F_{t_1,t_2,...,t_n}(\mathrm{B}_{\alpha^{-1}(1)}, \mathrm{B}_{\alpha^{-1}(2)},..., \mathrm{B}_{\alpha^{-1}(\mathrm{n})}). \end{aligned}$$

2. Verification of the completeness condition (4)

$$
\begin{aligned}
&F_{t_1,t_2},...,t_n,...,t_{n+m}(\mathrm{B}_1,\mathrm{B}_2,...,\mathrm{B}_\mathrm{n},\mathrm{R},...,\mathrm{R}) = \\
&= \mathrm{E}^\mathrm{Y}\left[\begin{array}{l} \mathrm{P}^\mathrm{X}\left(\begin{array}{l} X_{t_1}^{x_1} \in B_1, X_{t_1}^{x_2} \in B_2, ..., X_{t_n}^{x_n} \in B_n, \\ X_{t_{n+1}}^{x_{n+1}} \in R, ..., X_{t_{n+m}}^{x_{n+m}} \in R \end{array}\right)\Big| \\ \Big| \begin{array}{l} x_1 = Y_{t_1}^z, x_2 = Y_{t_2}^z, ..., x_n = Y_{t_n}^z, \\ x_{n+1} = Y_{t_{n+1}}^z, ..., x_{m+n} = Y_{t_{m+n}}^z \end{array} \end{array}\right] = \\
&= \mathrm{E}^\mathrm{Y}\left[\begin{array}{l} \mathrm{P}^\mathrm{X}(X_{t_1}^{x_1} \in B_1, X_{t_1}^{x_2} \in B_2, ..., X_{t_n}^{x_n} \in B_n)\big| \\ \big|x_1 = Y_{t_1}^z, x_2 = Y_{t_2}^z, ..., x_n = Y_{t_n}^z \end{array}\right] = \\
&= F_{t_1,t_2,...,t_n}(\mathrm{B}_1,\mathrm{B}_2,...,\mathrm{B}_\mathrm{n})
\end{aligned}
$$

for an arbitrary m.

Using this theorem we can give the next definition:

Definition 1.1. *We call an Ito-Levi process the stochastic process Z_t with finite-dimensional distributions* (2)*. We denote its probability space by $\left(\Omega^Z, \mathcal{F}_t^Z, P^Z\right)$.*

Analogously we can define a multidimensional Ito-Levy process, but for clearness we shall study only the one-dimensional case.

2 Construction

At this point we shall construct the Ito-Levy process and its probability space. We shall show that the Ito and the Levy processes are particular cases of Ito-Levy processes. Let $\Omega = \Omega^X \times \Omega^Y$. For all Borel B, define

$$A_t^{B,z} = \bigcup_{x \in R} \left\{(\omega^X, \omega^Y) : Y_t^z(\omega^Y) = x, X_t^x(\omega^x) \in B\right\}.$$

For arbitrary moments $t_1, t_2, ..., t_n$ and Borel sets $B_1, B_2, ..., B_n$ let us define A_t^B as the section of the corresponding $A_{t_i}^{B_i,z}$, $i = 1, ..., n$. All choices of n, $t_1, t_2, ..., t_n < t$ and $B_1, B_2, ..., B_n$ define the σ-algebra $\mathcal{F}_t^z$, generated by Z_t^z. Analogously me can define σ-algebra $\mathcal{F}_t$, generated by Z_t^z for all z. We can define the probability measure with the finite-dimensional distribution examined in the previous section:

$$
\begin{aligned}
&P(Z_{t_1}^z \in \mathrm{B}_1, Z_{t_2}^z \in \mathrm{B}_2, ..., Z_{t_n}^z \in \mathrm{B}_\mathrm{n}) = \\
&\qquad = \mathrm{E}^\mathrm{Y}\left[\begin{array}{l} \mathrm{P}^\mathrm{X}(X_{t_1}^{x_1} \in B_1, X_{t_1}^{x_2} \in B_2, ..., X_{t_n}^{x_n} \in B_n)\big| \\ \big|x_1 = Y_{t_1}^z, x_2 = Y_{t_2}^z, ..., x_n = Y_{t_n}^z \end{array}\right].
\end{aligned}
$$

In this way we have defined the probability space of the process Z. Now we shall construct the process. Let us have a look at the process

$$S_t^z(\omega) = S_t^z(\omega^X, \omega^Y) = X_t^x(\omega^X)\big|\, x = Y_t^z(\omega^Y).$$

We shall show that this process has the same distribution as Z_t^z. From the independence of X_t and Y_t we obtain:

$$\begin{aligned}
&\mathrm{P}^{\mathrm{X}}(S_{t_1}^z \in B_1, S_{t_1}^z \in B_2, ..., S_{t_n}^z \in B_n) = \\
&\qquad = \int\limits_{\mathbf{R}^n} \begin{array}{l} \mathrm{P}^{\mathrm{X}}(X_{t_1}^{x_1} \in B_1, X_{t_2}^{x_2} \in B_2, ..., X_{t_n}^{x_n} \in B_n) \\ dP^{Y}\left(Y_{t_1}^z < x_1, Y_{t_2}^z < x_2, ..., Y_{t_n}^z < x_n\right) \end{array} = \\
&\qquad = \mathrm{E}^{\mathrm{Y}}\left[\begin{array}{l} \mathrm{P}^{\mathrm{X}}(X_{t_1}^{x_1} \in B_1, X_{t_2}^{x_2} \in B_2, ..., X_{t_n}^{x_n} \in B_n)\big| \\ \big| x_1 = Y_{t_1}^z, x_2 = Y_{t_2}^z, ..., x_n = Y_{t_n}^z \end{array} \right],
\end{aligned}$$

which is exactly the distribution of Z_t^z.

Proposition 2.1. *The Levy processes are a particular case of the Ito-Levy processes.*

Proof. Let Z be a Levy process with generating triple $(\gamma, A, \Pi(dx))$. Using the *Levy-Ito decomposition theorem* (See [9, theorem 19.2.]) we can deduce that there exist two independent processes Z^1 and Z^2 with corresponding generating triples $(\gamma, A, 0)$ and $(0, 0, \Pi(dx))$, which satisfy $Z = Z^1 + Z^2$. Therefore

$$\begin{aligned}
&P(Z_{t_1} \in B_1, Z_{t_2} \in B_2, ..., Z_{t_n} \in B_n) = \\
&\qquad = P(Z_{t_1}^1 + Z_{t_1}^2 \in B_1, Z_{t_2}^1 + Z_{t_2}^2 \in B_2, ..., Z_{t_n}^1 + Z_{t_n}^2 \in B_n) = \\
&\qquad = \int\limits_{\mathbf{R}^n} \begin{array}{l} P(Z_{t_1}^1 + y_1 \in B_1, Z_{t_2}^1 + y_2 \in B_2, ..., Z_{t_n}^1 + y_n \in B_n) \\ dP(Z_{t_1}^2 < y_1, Z_{t_2}^2 < y_2, ..., Z_{t_n}^2 < y_n) \end{array} = \\
&\qquad = \mathrm{E}^{\mathrm{Z}^2}\left[\begin{array}{l} P(Z_{t_1}^1 + y_1 \in B_1, Z_{t_2}^1 + y_2 \in B_2, ..., Z_{t_n}^1 + y_n \in B_n)\big| \\ \big| y_1 = Z_{t_1}^2, y_2 = Z_{t_2}^2, ..., y_n = Z_{t_n}^2) \end{array} \right].
\end{aligned}$$

The process Z^1 is Gaussian, i.e. a particular case of the Ito processes (the initial condition appears as the sum) and Z^2 is purely non-Gaussian. This shows that Z is a particular case of the Ito-Levy processes. □

3 Expectation

Now we shall deduce the formula for the expectation of a function of Ito-Levi processes.

Proposition 3.1. $E\left(h(Z_t^z)\right) = E^Y\left[E^X\left(h(X_t^x)\right)\big| x = Y_t^z\right].$

Proof.

$$\begin{aligned}
E\left(h(Z_t^z)\right) &= \int\limits_{-\infty}^{+\infty} h(y) dP(Z_t^z < y) = \\
&= \int\limits_{-\infty}^{+\infty} h(y) \frac{d}{dy} E^Y\left[P^X(X_t^x < y)\big| x = Y_t^z\right] =
\end{aligned}$$

$$= \int_{-\infty}^{+\infty} h(y) \frac{d}{dy} \int_{-\infty}^{+\infty} P^X(X_t^x < y) dP^Y(Y_t^z < x) dy =$$
$$= \int_{-\infty}^{+\infty} h(y) \int_{-\infty}^{+\infty} \frac{d}{dy} P^X(X_t^x < y) dP^Y(Y_t^z < x) dy =$$
$$= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h(y) \frac{d}{dy} P^X(X_t^x < y) dy dP^Y(Y_t^z < x) =$$
$$= \int_{-\infty}^{+\infty} E^X [h(X_t^x)] \, dP^Y(Y_t^z < x) =$$
$$= E^Y \left[E^X [h(X_t^x)] \middle| \, x = Y_t^z \right].$$

□

4 Conditional Expectation

In this section we shall deduce the formula for the conditional expectation of functions of Ito-Levi processes.

Proposition 4.1.

$$E\left[g(Z_t^z)/\mathcal{F}_s\right]\left(\omega^X, \omega^Y\right) =$$
$$= E^Y \left[E^X \left[g\left(X_{t-s}^\alpha\right)\right] \middle| \, \alpha = X_s^x(\omega^X) \middle| \, x = Y_{t-s}^\beta \right] \middle| \, \beta = Y_s^z(\omega^Y).$$

Proof. Let $B \in \mathcal{F}_s$ and put

$$B^X(x) = (X_s^x)^{-1}\left(Z_s^z(B)\right),$$
$$B^Y(x) = \left\{\omega^Y : Y_s^z = x\right\}.$$

We denote that $B^X(x) \in \mathcal{F}_s^X$ and $B^Y(x) \in \mathcal{F}_s^Y$. Thus

$$\int_{\omega \in B} g\left(Z_t^z(\omega)\right) P(d\omega) =$$
$$= \int_{-\infty}^{+\infty} \int_{\omega^Y \in B^Y(x)} \left[\int_{\omega^X \in B^X(x)} g\left(X_t^\gamma(\omega^X)\right) \middle| \, \gamma = Y_t^z(\omega^Y) P^X(d\omega^X) \right] P^Y(d\omega^Y) dx =$$
$$= \int_{-\infty}^{+\infty} \int_{\omega^Y \in B^Y(x)} \left[\int_{\omega^X \in B^X(x)} g\left(X_t^\gamma(\omega^X)\right) P^X(d\omega^X) \right] \middle| \, \gamma = Y_t^z(\omega^Y) P^Y(d\omega^Y) dx. \tag{5}$$

Using the Markov property of X

$$E^X\left(g(X_t^\gamma)/\mathcal{F}_s^X\right) = E^X\left(g(X_{t-s}^\alpha)\right)\big|\,\alpha = X_s^\gamma,$$

and therefore

$$\begin{aligned}\int\limits_{\omega^X \in B^X(x)} & g\left(X_t^\gamma(\omega^X)\right) P^X(d\omega^X) = \\ &= \int\limits_{\omega^X \in B^X(x)} E^X\left[g\left(X_{t-s}^\alpha\right)\right]\big|\,\alpha = X_s^\gamma(\omega^X) P^X(d\omega^X).\end{aligned} \tag{6}$$

After replacing (6) in (5) we obtain:

$$\begin{aligned}&\int\limits_{\omega \in B} g\left(Z_t^z(\omega)\right) P(d\omega) = \\ &= \int\limits_{-\infty}^{+\infty} \int\limits_{\omega^Y \in B^Y(x)} \left[\int\limits_{\omega^X \in B^X(x)} g\left(X_t^\gamma(\omega^X)\right) P^X(d\omega^X)\right]\Bigg|\,\gamma = Y_t^z(\omega^Y) P^Y(d\omega^Y) dx = \\ &= \int\limits_{-\infty}^{+\infty} \int\limits_{\omega^Y \in B^Y(x)} \left[\left(\int\limits_{\omega^X \in B^X(x)} \begin{pmatrix} E^X\left[g\left(X_{t-s}^\alpha\right)\right]\big| \\ \big|\alpha = X_s^\gamma(\omega^X) \end{pmatrix} P^X(d\omega^X)\right)\Bigg| \atop \big|\gamma = Y_t^z(\omega^Y)\right] P^Y(d\omega^Y) dx.\end{aligned} \tag{7}$$

Using the Markov property of Y

$$E^Y\left(h(Y_t^z)/\mathcal{F}_s^Y\right) = E^Y\left(h(Y_{t-s}^\alpha)\right)\big|\,\alpha = Y_s^z$$

for

$$h(\gamma) = \int\limits_{\omega^X \in B^X(x)} E^X\left[g\left(X_{t-s}^\alpha\right)\right]\big|\,\alpha = X_s^\gamma(\omega^X) P^X(d\omega^X),$$

and therefore

$$\begin{aligned}\int\limits_{\omega^Y \in B^Y(x)} & h\left(Y_t^z(\omega^Y)\right) P^Y(d\omega^Y) = \\ &= \int\limits_{\omega^Y \in B^Y(x)} E^Y\left[h\left(Y_{t-s}^\alpha\right)\right]\big|\,\alpha = Y_s^z(\omega^Y) P^Y(d\omega^Y).\end{aligned} \tag{8}$$

After replacing (8) in (7) we obtain:

$$\int\limits_{\omega\in B} g\left(Z_t^z(\omega)\right)P(d\omega)=$$

$$=\int\limits_{-\infty}^{+\infty}\int\limits_{\omega^Y\in B^Y(x)}\left[\left(\int\limits_{\omega^X\in B^X(x)}\begin{pmatrix}E^X\left[g\left(X_{t-s}^{\alpha}\right)\right]\big|\\ \big|\alpha=X_s^{\gamma}(\omega^X)\end{pmatrix}P^X(d\omega^X)\right)\middle|\,\middle| \gamma=Y_t^z(\omega^Y)\right]P^Y(d\omega^Y)dx=$$

$$=\int\limits_{-\infty}^{+\infty}\int\limits_{\omega^Y\in B^Y(x)}\left[E^Y\left[\int\limits_{\omega^X\in B^X(x)}\begin{pmatrix}E^X\left[g\left(X_{t-s}^{\alpha}\right)\right]\big|\\ \big|\alpha=X_s^{\gamma}(\omega^X)\end{pmatrix}P^X(d\omega^X)\middle|\,\middle|\gamma=Y_{t-s}^{\beta}\right]\middle|\,\middle|\beta=Y_s^z(\omega^Y)\right]P^Y(d\omega^Y)dx.$$

We change the places of the expectation and the integration in

$$E^Y\left[\int\limits_{\omega^X\in B^X(x)}E^X\left[g\left(X_{t-s}^{\alpha}\right)\right]\big|\,\alpha=X_s^{\gamma}(\omega^X)P^X(d\omega^X)\middle|\,\gamma=Y_{t-s}^{\beta}\right]$$

to obtain

$$\int\limits_{\omega\in B} g\left(Z_t^z(\omega)\right)P(d\omega)=$$

$$=\int\limits_{-\infty}^{+\infty}\int\limits_{\omega^Y\in B^Y(x)}\left[E^Y\left[\int\limits_{\omega^X\in B^X(x)}\begin{pmatrix}E^X\left[g\left(X_{t-s}^{\alpha}\right)\right]\big|\\ \big|\alpha=X_s^{\gamma}(\omega^X)\end{pmatrix}P^X(d\omega^X)\middle|\,\middle|\gamma=Y_{t-s}^{\beta}\right]\middle|\,\middle|\beta=Y_z^z(\omega^Y)\right]P^Y(d\omega^Y)dx=$$

$$=\int\limits_{-\infty}^{+\infty}\int\limits_{\omega^Y\in B^Y(x)}\left[\int\limits_{\omega^X\in B^X(x)}E^Y\begin{bmatrix}E^X\left[g\left(X_{t-s}^{\alpha}\right)\right]\big|\\ \big|\,\alpha=X_s^{\gamma}(\omega^X)\big|\\ \big|\gamma=Y_{t-s}^{\beta}\end{bmatrix}P^X(d\omega^X)\middle|\,\middle|\beta=Y_s^z(\omega^Y)\right]P^Y(d\omega^Y)dx=$$

$$=\int\limits_{-\infty}^{+\infty}\int\limits_{\omega^Y\in B^Y(x)}\int\limits_{\omega^X\in B^X(x)}\left[E^Y\begin{bmatrix}E^X\left[g\left(X_{t-s}^{\alpha}\right)\right]\big|\\ \big|\,\alpha=X_s^{\gamma}(\omega^X)\big|\\ \big|\gamma=Y_{t-s}^{\beta}\end{bmatrix}\middle|\,\middle|\beta=Y_s^z(\omega^Y)\right]P^X(d\omega^X)P^Y(d\omega^Y)dx=$$

$$= \int\limits_{\omega \in B} \left[E^Y \left[\begin{array}{l} E^X \left[g\left(X^{\alpha}_{t-s} \right) \right] \big| \\ \big| \alpha = X^{\gamma}_s(\omega^X) \big| \\ \big| \gamma = Y^{\beta}_{t-s} \end{array} \right] \Bigg| \beta = Y^z_s(\omega^Y) \right] P(d\omega).$$

We must show only that

$$E^Y \left[E^X \left[g\left(X^{\alpha}_{t-s} \right) \right] \big| \, \alpha = X^x_s(\omega^X) \big| \, x = Y^{\beta}_{t-s} \right] \Big| \, \beta = Y^z_s(\omega^Y)$$

is $\mathcal{F}_s$-adapted. We have

$$\begin{aligned} & E^Y \left[E^X \left[g\left(X^{\alpha}_{t-s} \right) \right] \big| \, \alpha = X^x_s(\omega^X) \big| \, x = Y^{\beta}_{t-s} \right] \Big| \, \beta = Y^z_s(\omega^Y) = \\ & = E^Y \left[E^X \left[g\left(X^{\alpha}_{t-s} \right) \right] \big| \, \alpha = X^x_s(\omega^X) \big| \, x = Y_{t-s} + Y^z_s(\omega^Y) \right] = \\ & = E^Y \left[E^X \left[g\left(X^{\alpha}_{t-s} \right) \right] \big| \, \alpha = X^x_s(\omega^X) \big| \, x = Y^{\beta}_s(\omega^Y) \big| \, \beta = Y^z_{t-s} \right] = \\ & = E^Y \left[E^X \left[g\left(X^{\alpha}_{t-s} \right) \right] \big| \, \alpha = Z^{\beta}_s(\omega) \big| \, \beta = Y^z_{t-s} \right]. \end{aligned}$$

We see from here that

$$E^Y \left[E^X \left[g\left(X^{\alpha}_{t-s} \right) \right] \big| \, \alpha = X^x_s(\omega^X) \big| \, x = Y^{\beta}_{t-s} \right] \Big| \, \beta = Y^z_s(\omega^Y)$$

is a consequence of Z^{β}_s for all β. Therefore it is adapted to the σ-algebra $\mathcal{F}_s$, generated by Z^{β}_s for all β. So

$$E\left[g(Z^z_t)/\mathcal{F}_s \right] = E^Y \left[E^X \left[g\left(X^{\alpha}_{t-s} \right) \right] \big| \, \alpha = X^x_s(\omega^X) \big| \, x = Y^{\beta}_{t-s} \right] \Big| \, \beta = Y^z_s(\omega^Y).$$

□

Let $\mathcal{F}^z$ be the σ-algebra generated by the process Z^z_t for a particular z and $\mathcal{H}$ be the σ-algebra generated by the processes X^x_t for all x, and Y.

Analogously we can show that the conditional expectation w.r.t. the σ-algebra $\mathcal{H}$ is again

$$E\left[g(Z^z_t)/\mathcal{H}_s \right] = E^Y \left[E^X \left[g\left(X^{\alpha}_{t-s} \right) \right] \big| \, \alpha = X^x_s(\omega^X) \big| \, x = Y^{\beta}_{t-s} \right] \Big| \, \beta = Y^z_s(\omega^Y).$$

But for the conditional expectation w.r.t. the σ-algebra $\mathcal{F}^z$ we can only say that

$$\begin{aligned} & \int\limits_{\omega \in B} E\left[g(Z^z_t) \right] P\left(d\omega \right) = \\ & = \int\limits_{\omega \in B} E^Y \left[E^X \left[g\left(X^{\alpha}_{t-s} \right) \right] \big| \, \alpha = X^x_s(\omega^X) \big| \, x = Y^{\beta}_{t-s} \right] \Big| \, \beta = Y^z_s(\omega^Y) P\left(d\omega \right) \end{aligned}$$

because $E^Y\left[E^X\left[g\left(X^\alpha_{t-s}\right)\right]\big|\,\alpha=X^x_s(\omega^X)\big|\,x=Y^\beta_{t-s}\right]\Big|\,\beta=Y^z_s(\omega^Y)$ is not measurable w.r.t. the σ-algebra $\mathcal{F}^z$.

In the same way we can prove the analogous proposition for the conditional expectation of the "discounting" process.

Proposition 4.2.

$$E\left[\frac{g\left(Z^z_t\right)}{\exp\left(\int\limits_s^t a\left(\overline{Z}^{\overline{z}}_u\right)ds\right)}\middle|\mathcal{F}_s\right]\left(\omega^X,\omega^Y\right)=$$

$$=E^Y\left[E^X\left[\frac{g\left(X^\alpha_{t-s}\right)\big|\,\alpha=X^x_s\left(\omega^X\right)\big|\,x=Y^y_{t-s}}{\exp\left(\int\limits_0^{t-s}a\left(\overline{X}^\alpha_u\right)\Big|\,\alpha=\overline{X}^x_s\left(\omega^X\right)\Big|\,x=\overline{Y}^{\overline{y}}_u du\right)}\right]\right]\Bigg|$$
$$\Big|y=Y^z_s\left(\omega^Y\right);\overline{y}=\overline{Y}^z_s\left(\omega^Y\right).$$

5 Markov Property

We can see from the form of the conditional expectation that the Ito-Levi processes are not markovian at all. The reason for this is that the defining of the markovian semi group for the Ito and the Levi processes is done in a different way. While for the Ito processes the markovian semi group is $Pg\left(x\right)=E\left[g\left(X^x_t\right)\right]$, where the variable x is the initial condition for the stochastic differential equation (1), for the Levy processes we have for the markovian semi group

$$Pg\left(y\right)=E\left[g\left(Y^y_t\right)\right]=E\left[g\left(Y_t+y\right)\right],$$

i.e. the variable y is added to the process Y. This difference shows why we can not exchange the places of X and Y in $X^x_s(\omega^X)\big|\,x=Y^\beta_{t-s}$. So we can say that if the form of X^x_t permits us to change the places of X and Y, then the process Z_t is markovian.

Example 5.1. Let

$$\begin{aligned}dX^x_t&=\mu dt+\sigma dB_t\\ X^x_0&=x,\quad x\in R^n,\end{aligned}$$

i.e. X^x_t is the ordinary Gaussian process $x+\mu t+\sigma B_t$. Then

$$\begin{aligned}X^x_s(\omega^X)\big|\,x=Y^\beta_{t-s}&=\\ &=x+\mu t+\sigma B_t|\,x=Y_{t-s}+\beta=\\ &=Y_{t-s}+x|\,x=\beta+\mu t+\sigma B_t\end{aligned}$$

and therefore we have for the conditional expectation

$$\begin{aligned} E\left[g(Z_t^z)/\mathcal{F}_s\right] = & \\ &= E^Y\left[\, E^X\left[g\left(X_{t-s}^{\alpha}\right)\right]\Big|\,\alpha = X_s^x(\omega^X)\Big|\,x = Y_{t-s}^{\beta}\right]\Big|\,\beta = Y_s^z(\omega^Y) = \\ &= E^Y\left[\, E^X\left[g\left(X_{t-s}^{\alpha}\right)\right]\Big|\,\alpha = Y_{t-s}^{x}\Big|\,x = X_s^{\beta}\right]\Big|\,\beta = Y_s^z(\omega^Y) = \\ &= E\left[g(Z_{t-s}^x)\right]\Big|\,x = Z_s^z, \end{aligned}$$

which shows the Markov property in this example.

On the other hand if we generalize the idea of the Markov property, we can give some more precise results. The Markov property is closely related to the history of the process. In this situation if we consider the process from the point of view of the history of a more general class of processes, then the Ito-Levy processes turn out markovian. Mathematically these arguments look in the following way:

Let again $\mathcal{F}^z$ be the σ-algebra generated by the process Z_t^z for a particular z; $\mathcal{F}$ be the σ-algebra generated by the processes Z_t^z for all z; and at the end, $\mathcal{H}$ be the σ-algebra generated by the processes X_t^x for all x, and Y. Note that $\mathcal{F}^z \subset \mathcal{F} \subset \mathcal{H}$. Under these circumstance the Ito-Levi process Z_t^z is markovian w.r.t. the σ-algebras $\mathcal{F}$ and $\mathcal{H}$, but not w.r.t. its own σ-algebra $\mathcal{F}^z$.

6 Infinitesimal Generator

In this section we shall find the infinitesimal generator of an Ito-Levy process. At the beginning let us observe the case when t=0.

Proposition 6.1. *The infinitesimal generator of an Ito-Levy process for t=0 is:*

$$\begin{aligned} Ag(z) = &\left(\mu(z)\frac{\partial}{\partial z}g(z) + \frac{1}{2}\sigma^2(z)\frac{\partial^2}{\partial z^2}g(z)\right) + \\ &+ \int_{-\infty}^{+\infty} g(z+y) - g(z) - I_D(y)y\frac{\partial}{\partial z}g(z)\Pi(dy) + b\frac{\partial}{\partial z}g\left(z\right) \end{aligned}$$

for every bounded and continuous function g.

Proof.

$$\begin{aligned} Ag(z) &= \lim_{t\to 0}\frac{E\left[g(Z_t^z)\right] - g(z)}{t} = \\ &= \lim_{t\to 0}\frac{E^Y\left[E^X\left[g(X_t^x)\right]\big|\,x = Y_t^z\right] - g(z)}{t} = \end{aligned}$$

$$
\begin{aligned}
&= \lim_{t\to 0} \frac{E^Y\left[E^X\left[g(X_t^x)\right]\middle| x = Y_t^z\right] - g(z) \pm E^Y\left[g(Y_t^z)\right]}{t} = \\
&= \lim_{t\to 0} \frac{E^Y\left[E^X\left[g(X_t^x)\right]\middle| x = Y_t^z\right] - E^Y\left[g(Y_t^z)\right]}{t} + \\
&+ \lim_{t\to 0} \frac{E^Y\left[g(Y_t^z)\right] - g(z)}{t} = \\
&= L_1 + L_2. \qquad (9)
\end{aligned}
$$

We shall study separately the two limits L_1 and L_2:

$$
\begin{aligned}
L_1 &= \lim_{t\to 0} \frac{E^Y\left[E^X\left[g(X_t^x)\right]\middle| x = Y_t^z\right] - E^Y\left[g(Y_t^z)\right]}{t} = \\
&= \lim_{t\to 0} \frac{E^Y\left[E^X\left[g(X_t^x)\right]\middle| x = Y_t^z - g(Y_t^z)\right]}{t} = \\
&= \lim_{t\to 0} \frac{E^Y\left[E^X\left[g(X_t^x) - g(x)\right]\middle| x = Y_t^z\right]}{t} = \\
&= \lim_{t\to 0} \frac{E^Y\left[E^X\left[g(X_t^x) - g(x)\right]\middle| x = Y_t^z\right] \pm E^X\left[g(X_t^z) - g(z)\right]}{t} = \\
&= \lim_{t\to 0} \frac{E^Y\left[E^X\left[g(X_t^x) - g(x)\right]\middle| x = Y_t^z\right] - E^X\left[g(X_t^z) - g(z)\right]}{t} + \\
&+ \lim_{t\to 0} \frac{E^X\left[g(X_t^z) - g(z)\right]}{t} = \\
&= A^Y E^X\left[g(X_0^z) - g(z)\right] + A^X g\left(z\right) = \\
&= \mu\left(z\right)\frac{\partial}{\partial z} g(z) + \frac{1}{2}\sigma^2\left(z\right)\frac{\partial^2}{\partial z^2} g(z)
\end{aligned}
$$

and

$$
L_2 = \int_{-\infty}^{+\infty} g(z+y) - g(z) - I_D(y) y \frac{\partial}{\partial z} g(z) \Pi(dy) + b\frac{\partial}{\partial z} g\left(z\right)
$$

and therefore we get for the infinitesimal generator (9):

$$
\begin{aligned}
Ag(z) &= L_1 + L_2 = \\
&= \left(\mu(z)\frac{\partial}{\partial z} g(z) + \frac{1}{2}\sigma^2(z)\frac{\partial^2}{\partial z^2} g(z)\right) + \\
&+ \int_{-\infty}^{+\infty} g(z+y) - g(z) - I_D(y) y \frac{\partial}{\partial z} g(z) \Pi(dy) + b\frac{\partial}{\partial z} g\left(z\right).
\end{aligned}
$$

$\square$

Note that this infinitesimal generator is the sum of the infinitesimal generators of X_t and Y_t, i.e.

$$Ag\left(z\right) = A^X g\left(z\right) + A^Y g\left(z\right).$$

Now we shall study the case when $s > 0$.

Proposition 6.2. *The infinitesimal generator of an Ito-Levy process for an arbitrary moment s is:*

$$\begin{aligned} A_s g\left(z\right) = &E^Y\left[E^X\left\{A^X g(\alpha)\middle|\, \alpha = X_s^x\right\}\middle|\, x = Y_s^z\right] + \\ &+ E^Y\left\{A^Y E^X\left[g(X_s^\alpha)\right]\middle|\, \alpha = Y_s^z\right\}. \end{aligned}$$

Proof. We shall only sketch out the proof, because it is closely related to the proof of the previous proposition.

$$\begin{aligned} A_s g(z) &= \lim_{t\to 0}\frac{E\left[g(Z_{s+t}^z)\right] - E\left[g(Z_s^z)\right]}{t} = \\ &= \lim_{t\to 0}\frac{E^Y\left[E^X\left[g(X_{s+t}^x)\right]\middle|\, Y_{s+t}^z\right] - E^Y\left[E^X\left[g(X_s^x)\right]\middle|\, x = Y_s^z\right]}{t} = \\ &= \lim_{t\to 0}\frac{E^Y\left[E^X g\left[(X_{s+t}^x)\right]\middle|\, x = Y_{s+t}^z\right] - E^Y\left[E^X\left[g(X_s^x)\right]\middle|\, x = Y_s^z\right]}{t} \pm \\ &\pm \lim_{t\to 0}\frac{E^Y\left[E^X\left[g(X_s^x)\right]\middle|\, x = Y_{s+t}^z\right]}{t} = \\ &= \lim_{t\to 0}\frac{E^Y\left[E^X\left[g(X_{s+t}^x)\right]\middle|\, x = Y_{s+t}^z\right] - E^Y\left[E^X\left[g(X_s^x)\right]\middle|\, x = Y_{s+t}^z\right]}{t} + \\ &+ \lim_{t\to 0}\frac{E^Y\left[E^X\left[g(X_s^x)\right]\middle|\, x = Y_{s+t}^z\right] - E^Y\left[E^X\left[g(X_s^x)\right]\middle|\, x = Y_s^z\right]}{t} = \\ &= L_1 + L_2. \end{aligned}$$

We shall study again separately the two limits L_1 and L_2:

$$\begin{aligned} L_1 &= \lim_{t\to 0}\frac{E^Y\left[E^X\left[g(X_{s+t}^x)\right]\middle|\, x = Y_{s+t}^z\right] - E^Y\left[E^X\left[g(X_s^x)\right]\middle|\, x = Y_{s+t}^z\right]}{t} = \\ &= \lim_{t\to 0}\frac{E^Y\left[E^X\left[g(X_{s+t}^x) - g(X_s^x)\right]\middle|\, x = Y_{s+t}^z\right]}{t} = \\ &= \lim_{t\to 0}\frac{E^Y\left[E^X\left\{E^X\left[g(X_t^x) - g(x)\right]\middle|\, \alpha = X_s^x\right\}\middle|\, x = Y_{s+t}^z\right]}{t} = \\ &= E^Y\left[E^X\left\{A^X g(\alpha)\middle|\, \alpha = X_s^x\right\}\middle|\, x = Y_s^z\right] \end{aligned}$$

and

$$
\begin{aligned}
L_2 &= \lim_{t\to 0} \frac{E^Y\left[E^X\left[g(X_s^x)\right]\middle| x = Y_{s+t}^z\right] - E^Y\left[E^X\left[g(X_s^x)\right]\middle| x = Y_s^z\right]}{t} = \\
&= \lim_{t\to 0} \frac{E^Y\left[E^X\left[g(X_s^x)\right]\middle| x = Y_{s+t}^z - E^X\left[g(X_s^x)\right]\middle| x = Y_s^z\right]}{t} = \\
&= \lim_{t\to 0} \frac{E^Y\left\{E^Y\left[E^X\left[g(X_s^x)\right]\middle| x = Y_t^\alpha - E^X\left[g(X_s^\alpha)\right]\right]\middle| \alpha = Y_s^z\right\}}{t} = \\
&= E^Y\left\{A^Y E^X\left[g(X_s^\alpha)\right]\middle|\alpha = Y_s^z\right\}
\end{aligned}
$$

or finaly

$$
\begin{aligned}
A_s g\left(z\right) = E^Y & \left[E^X\left\{A^X g(\alpha)\middle|\, \alpha = X_s^x\right\}\middle|\, x = Y_s^z\right] + \\
&+ E^Y\left\{A^Y E^X\left[g(X_s^\alpha)\right]\middle|\, \alpha = Y_s^z\right\}.
\end{aligned}
$$

□

7 Uniqueness of the presentation

To guarantee the uniqueness of the presentation of an Ito-Levy process we shall require the next condition:

Condition 7.1. *If the functions μ and σ, which are the characteristics of the Ito process, are constants (i.e. an ordinary Gaussian process), then $b = 0$.*

We shall show that under this condition, the presentation of an Ito-Levy process is unique.

Proposition 7.1. *Suppose that $X_t^x|\, x = Y_t^z \equiv \left.\overline{X}_t^x\right|\, x = Y_t^z$ and the previous condition is satisfied. Then*

$$
\begin{aligned}
\mu\left(x\right) &\equiv \overline{\mu}\left(x\right) \\
\sigma\left(x\right) &= \overline{\sigma}\left(x\right) \\
b &= \overline{b} \\
\Pi &\equiv \overline{\Pi}.
\end{aligned}
$$

Proof. Let us look at the infinitesimal generators:

$$
\begin{aligned}
Ag(z) = & \left(\mu(z)\frac{\partial}{\partial z}g(z) + \frac{1}{2}\sigma^2(z)\frac{\partial^2}{\partial z^2}g(z)\right) + \\
&+ b\frac{\partial}{\partial z}g(z) + \int_{-\infty}^{+\infty} g(z+y) - g(z) - I_D(y)y\frac{\partial}{\partial z}g(z)\Pi(dy) =
\end{aligned}
$$

$$= (\mu(z) + b)\,\frac{\partial}{\partial z} g(z) + \frac{1}{2}\sigma^2(z)\frac{\partial^2}{\partial z^2} g(z) +$$
$$+ \int_{-\infty}^{+\infty} g(z+y) - g(z) - I_D(y) y \frac{\partial}{\partial z} g(z)\Pi(dy)$$

$$\overline{A}g(z) = \left(\overline{\mu}(z)\frac{\partial}{\partial z}g(z) + \frac{1}{2}\overline{\sigma}^2(z)\frac{\partial^2}{\partial z^2}g(z)\right) +$$
$$+ \overline{b}\frac{\partial}{\partial z}g(z) + \int_{-\infty}^{+\infty} g(z+y) - g(z) - I_D(y) y \frac{\partial}{\partial z} g(z)\overline{\Pi}(dy) =$$
$$= \left(\overline{\mu}(z) + \overline{b}\right)\frac{\partial}{\partial z}g(z) + \frac{1}{2}\overline{\sigma}^2(z)\frac{\partial^2}{\partial z^2}g(z) +$$
$$+ \int_{-\infty}^{+\infty} g(z+y) - g(z) - I_D(y) y \frac{\partial}{\partial z} g(z)\overline{\Pi}(dy).$$

Since the infinitesimal generators are equal, then $\Pi \equiv \overline{\Pi}$ (from the pseudo-differential part). Also if $Y_t = bt + U_t$ and $\overline{Y}_t = \overline{b}t + \overline{U}_t$ is the Ito-Levy decomposition, then $U_t = \overline{U}_t$.

Suppose that $b = \overline{b}$. Then we see that $\mu(x) \equiv \overline{\mu}(x)$ and $\sigma(x) = \overline{\sigma}(x)$ and therefore the proposition is proved.

Let $b \neq \overline{b}$. From the coincidence of the two infinitesimal generators we have:

$$\mu(x) - \overline{\mu}(x) + b - \overline{b} = 0$$
$$\sigma(x) = \overline{\sigma}(x).$$

On the other hand if we use *the Ito-Levy decomposition theorem* we obtain:

$$\left.\overline{X}_t^x\right|x = \overline{Y}_t^z \equiv X_t^x|\,x = Y_t^z =$$
$$= X_t^x|\,x = bt + U_t^z =$$
$$= X_t^x|\,x = \left(b - \overline{b}\right)t + \overline{b}t + U_t^z =$$
$$= X_t^x|\,x = \left(b - \overline{b}\right)t + y\big|\, y = \overline{b}t + \overline{U}_t^z$$

and therefore if $c = b - \overline{b}$, then $\overline{X}_t^y = X_t^x|\,x = ct + y$. If we look at the Ito process X_t^x as a function $X(t,x,A)$, where A is a multidimensional Ito process with characteristics α and β, then we have from the Ito formula (see [6, theorem 4.2.1]

$$dX(t,x,A) = \left[X_t(t,x,A) + \alpha X_A(t,x,A) + \frac{1}{2}tr\left(\beta\beta^T X_{AA}(t,x,A)\right)\right]dt +$$
$$+ \beta X_A(t,x,A)\,dB_t,$$

so

$$\mu\left(X\left(t,x,A\right)\right)=X_t\left(t,x,A\right)+\alpha X_A\left(t,x,A\right)+\frac{1}{2}tr\left(\beta\beta^T X_{AA}\left(t,x,A\right)\right)$$
$$\sigma\left(X\left(t,x,A\right)\right)=\beta X_A\left(t,x,A\right).$$

On the other hand $\overline{X}_t^y = X\left(t,y+ct,A\right)$ and therefore, using again the Ito formula, we have

$$\begin{aligned} d\overline{X}_t = & \begin{bmatrix} X_t\left(t,y+ct,A\right)+cX_x\left(t,y+ct,A\right)+\alpha X_A\left(t,y+ct,A\right)+ \\ +\frac{1}{2}tr\left(\beta\beta^T X_{AA}\left(t,y+ct,A\right)\right) \end{bmatrix} dt+ \\ & +\beta X_A\left(t,y+ct,A\right)dB_t = \\ & =\left[cX_x\left(t,y+ct,A\right)+\mu\left(\overline{X}_t\right)\right]dt+\sigma\left(\overline{X}_t\right)dB_t = \\ & =\left[cX_x\left(t,y+ct,A\right)-c+\overline{\mu}\left(\overline{X}_t\right)\right]dt+\sigma\left(\overline{X}_t\right)dB_t. \end{aligned}$$

Therefore $X_x = 1 \Rightarrow X_t^x = x + X_t^0$ and

$$X_t^x = X_t^0 + x = x + \int\limits_o^t \mu\left(X_s^0\right)ds + \int\limits_o^t \sigma\left(X_s^0\right)dB_s.$$

On the other hand

$$\begin{aligned} X_t^x &= x + \int\limits_o^t \mu\left(X_s^x\right)ds + \int\limits_o^t \sigma\left(X_s^x\right)dB_s = \\ &= x + \int\limits_o^t \mu\left(X_s^0 + x\right)ds + \int\limits_o^t \sigma\left(X_s^0 + x\right)dB_s. \end{aligned}$$

Since this is true for every arbitrary moment t, then μ and σ are constants. From the condition, which we imposed at the beginning of the section, we have that b and $\bar{b}$ are equal to zero. $\square$

8 Martingality

In this section we shall study the relationship between the martingality and the infinitesimal generator. We shall consider separately the cases when the process is an Ito process, a Levy process or an Ito-Levy process.

8.1 Martingality and the Ito processes

Let X_t^x be the Ito process (1). Then the next proposition is true:

Proposition 8.1. *The following statements are equivalent:*

1. *The process $g(X_t^x)$ is a martingale.*
2. *The process $g(X_t^x)$ has a constant expectation, i.e. $E[g(X_t^x)] = g(x)$.*
3. *$Ag(x) = 0$*

Proof. Since $g(X_t^x)$ is a martingale, then

$$E[g(X_t^x)] = E[g(X_0^x)] = g(x)$$

and therefore

$$Ag(x) = \lim_{t \to 0} \frac{E[g(X_t^x)] - g(x)}{t} = 0.$$

We must show only that point 1 follows from point 3. We shall use the Ito formula

$$g(X_t^x) = g(x) + \int_0^t Ag(\alpha)|\, \alpha = X_s^x ds + \int_0^t \sigma(X_s^x) g_x(X_s^x) dB_s =$$
$$= g(x) + \int_0^t \sigma(X_s^x) g_x(X_s^x) dB_s.$$

Since the drift is zero, then $g(X_t^x)$ is a martingale.

□

We shall prove the analogous proposition if $g(\cdot)$ is a function of the time and the process X_t^x.

Proposition 8.2. *The following statements are equivalent:*

1. *The process $g(t, X_t^x)$ is a martingale.*
2. *The process $g(t, X_t^x)$ has a constant expectation, i.e. $E[g(t, X_t^x)] = g(0, x)$.*
3. *$g_t(t, x) + Ag(t, x) = 0$, where $Ag(t, x) = \mu(x) g_x(t, x) + \frac{1}{2}\sigma^2(x) g_{xx}(t, x)$.*

Proof. We shall proof that point 3 follows from point 1, point 2 follows from point 3 and point 1 follows from point 2.

Let $g(t, X_t^x)$ be a martingale. Then the drift is zero, and using the Ito formula we state that

$$g_t\left(t, X_t^x\right)+\mu\left(X_t^x\right) g_x\left(t, X_t^x\right)+\frac{1}{2} \sigma^2\left(X_t^x\right) g_{xx}\left(t, X_t^x\right)=0$$

and therefore

$$g_t(t, x)+A g(t, x)=0.$$

Since the expectation of the stochastic integral is zero, then if $g_t(t, x)+A g(t, x)=0$, then

$$\begin{aligned}
E\left[g\left(t, X_t^x\right)\right]= & \, g(0, x)+ \\
& +E\left[\int_0^t g_t\left(s, X_s^x\right)+\mu\left(X_s^x\right) g_x\left(s, X_s^x\right)+\frac{\sigma^2\left(X_s^x\right)}{2} g_{xx}\left(s, X_s^x\right) ds\right]+ \\
& +E\left[\int_0^t \sigma\left(X_s^x\right) g_x\left(s, X_s^x\right) dB_s\right]= \\
& =g(0, x).
\end{aligned}$$

We must show only that if $E\left[g\left(t, X_t^x\right)\right]=g(0, x)$, then $g\left(t, X_t^x\right)$ is a martingale. We have

$$E\left[g\left(t+s, X_{t+s}^x\right) \mid \mathcal{F}_s\right]=E\left[g\left(t+s, X_t^\alpha\right)\right] \mid \alpha=X_s^x.$$

Take the $g\left(t+s, X_t^\alpha\right)$ in the Taylor's series

$$\begin{aligned}
E\left[g\left(t+s, X_t^\alpha\right)\right] & =E\left[\sum_{n=0}^{+\infty} \frac{s^n}{n!} g_t^{(n)}\left(t, X_t^\alpha\right)\right]= \\
& =\sum_{n=0}^{+\infty} \frac{s^n}{n!} E\left[g_t^{(n)}\left(t, X_t^\alpha\right)\right].
\end{aligned}$$

So we must consider the expectations of the derivatives of g. Let

$$u(t, x)=E\left[g\left(t, X_t^x\right)\right].$$

Since $E\left[g\left(t, X_t^x\right)\right]=g(0, x)$, then $u_t(t, x)=0$ and therefore

$$\begin{aligned}
0= & \, u_t(t, x)= \\
& =\lim_{\varepsilon \rightarrow 0} \frac{E\left[g\left(t+\varepsilon, X_{t+\varepsilon}^x\right)\right]-E\left[g\left(t, X_t^x\right)\right]}{\varepsilon}= \\
& =\lim_{\varepsilon \rightarrow 0} \frac{E\left[g\left(t+\varepsilon, X_{t+\varepsilon}^x\right)\right]-E\left[g\left(t, X_t^x\right)\right] \pm E\left[g\left(t, X_{t+\varepsilon}^x\right)\right]}{\varepsilon}=
\end{aligned}$$

$$
\begin{aligned}
&= \lim_{\varepsilon\to 0} \frac{E\left[g\left(t+\varepsilon, X_{t+\varepsilon}^x\right)\right] - E\left[g\left(t, X_{t+\varepsilon}^x\right)\right]}{\varepsilon} + \\
&+ \lim_{\varepsilon\to 0} \frac{E\left[g\left(t, X_{t+\varepsilon}^x\right)\right] - E\left[g\left(t, X_t^x\right)\right]}{\varepsilon} = \\
&= E\left[g_t\left(t, X_t^x\right)\right] + \lim_{\varepsilon\to 0} \frac{E\left[E\left[g\left(t, X_{t+\varepsilon}^x\right)/\mathcal{F}_\varepsilon\right]\right] - E\left[g\left(t, X_t^x\right)\right]}{\varepsilon} = \\
&= E\left[g_t\left(t, X_t^x\right)\right] + \lim_{\varepsilon\to 0} \frac{E\left[E\left[g\left(t, X_t^\alpha\right)_\varepsilon\right]\middle|\, \alpha = X_\varepsilon^x\right] - E\left[g\left(t, X_t^x\right)\right]}{\varepsilon} = \\
&= E\left[g_t\left(t, X_t^x\right)\right] + \lim_{\varepsilon\to 0} \frac{E\left[u\left(t, X_\varepsilon^x\right)\right] - u\left(t, x\right)}{\varepsilon} = \\
&= E\left[g_t\left(t, X_t^x\right)\right] + Au\left(t, x\right).
\end{aligned}
$$

Since $u(t,x)$ is independent of t variable, then $Au(t,x)$ is also independent of t. Therefore $E\left[g_t\left(t, X_t^x\right)\right]$ is independent of t as well. So $E\left[g_t\left(t, X_t^x\right)\right] = g_t(0,x)$. In fact we proved that if $E^Y\left[g\left(t, X_t^x\right)\right]$ is independent from t, then $E^Y\left[g_t\left(t, X_t^x\right)\right]$ is also independent of t. Analogously we can prove that $E^Y\left[g_t^{(n)}\left(t, X_t^x\right)\right]$ is independent of t and therefore $E^Y\left[g_t^{(n)}\left(t, X_t^x\right)\right] = g_t^{(n)}(0,x)$. Since

$$
E^Y\left[g\left(t+s, X_t^\alpha\right)\right] = \sum_{n=0}^{\infty} \frac{s^n}{n!} E^Y\left[g^{(n)}\left(t, X_t^\alpha\right)\right],
$$

then

$$
\begin{aligned}
E^Y\left[g\left(t+s, X_t^\alpha\right)\right] &= \sum_{n=0}^{\infty} \frac{s^n}{n!} E^Y\left[g^{(n)}\left(t, X_t^\alpha\right)\right] = \\
&= \sum_{n=0}^{\infty} \frac{s^n}{n!} g^{(n)}\left(0, \alpha\right) = \\
&= g\left(s, \alpha\right),
\end{aligned}
$$

from where we derive for the conditional expectation

$$
\begin{aligned}
E^Y\left[g\left(t+s, X_{t+s}^x\right)\middle|\,\mathcal{F}_s^Y\right] &= E^Y\left[g\left(t+s, X_t^\alpha\right)\right]\Big|\,\alpha = X_s^x \\
&= g\left(s, \alpha\right)\big|\,\alpha = X_s^x \\
&= g\left(s, X_s^x\right).
\end{aligned}
$$

□

8.2 Martingality and the Levy processes

We shall prove, that the analogous proposition is true if the process is of the Levy type. Let Y_t^y be a Levy process with the generating triple (b, σ, Π).

Proposition 8.3. *The following statements are equivalent:*

1. *The process $g(t, Y_t^y)$ is a martingale.*
2. *The process $g(t, Y_t^y)$ has a constant expectation,i.e. $E[g(t, Y_t^y)] = g(0, y)$.*
3. *$g_t(t, y) + Ag(t, y) = 0$, where*

$$\begin{aligned} Ag(t, y) = {} & bg_y(t, y) + \frac{1}{2}\sigma^2 g_{yy}(t, y) + \\ & + \int_{-\infty}^{+\infty} g(y+\alpha) - g(y) - I_D(\alpha)\alpha g_y(t, y)\Pi(d\alpha). \end{aligned}$$

Proof. We shall proof that point 3 follows from point 1, point 2 follows from point 3 and point 1 follows from point 2.

Let point 1 be true. Since $g(t, Y_t^y)$ is a martingale, then

$$\begin{aligned} g(t, Y_t^y) &= E\left[g\left(t+\varepsilon, Y_{t+\varepsilon}^y\right)\middle|\mathcal{F}_t\right] = \\ &= E\left[g\left(t+\varepsilon, Y_\varepsilon^\alpha\right)\right]\big|\,\alpha = Y_t^y. \end{aligned}$$

Therefore $g(t, \alpha) = E[g(t+\varepsilon, Y_\varepsilon^\alpha)]$. From here we obtain that

$$\begin{aligned} 0 &= \lim_{\varepsilon \to 0} \frac{E[g(t+\varepsilon, Y_\varepsilon^\alpha)] - g(t, \alpha)}{\varepsilon} = \\ &= \lim_{\varepsilon \to 0} \frac{E[g(t+\varepsilon, Y_\varepsilon^\alpha)] - g(t, \alpha) \pm E[g(t, Y_\varepsilon^\alpha)]}{\varepsilon} = \\ &= \lim_{\varepsilon \to 0} \frac{E[g(t+\varepsilon, Y_\varepsilon^\alpha)] - E[g(t, Y_\varepsilon^\alpha)]}{\varepsilon} + \lim_{\varepsilon \to 0} \frac{E[g(t, Y_\varepsilon^\alpha)] - g(t, \alpha)}{\varepsilon} = \\ &= g_t(t, \alpha) + Ag(t, \alpha). \end{aligned}$$

So we have that point 3 follows from point 1.

Let point 3 be true and $u(t, y) = E[g(t, Y_t^y)]$. We shall find the derivative w.r.t. t.

$$\begin{aligned} u_t(t, y) &= \lim_{\varepsilon \to 0} \frac{E\left[g\left(t+\varepsilon, Y_{t+\varepsilon}^y\right)\right] - E[g(t, Y_t^y)]}{\varepsilon} = \\ &= \lim_{\varepsilon \to 0} \frac{E\left[g\left(t+\varepsilon, Y_{t+\varepsilon}^y\right)\right] - E[g(t, Y_t^y)] \pm E\left[g\left(t, Y_{t+\varepsilon}^y\right)\right]}{\varepsilon} = \\ &= \lim_{\varepsilon \to 0} \frac{E\left[g\left(t+\varepsilon, Y_{t+\varepsilon}^y\right)\right] - E\left[g\left(t, Y_{t+\varepsilon}^y\right)\right]}{\varepsilon} + \\ &+ \lim_{\varepsilon \to 0} \frac{E\left[g\left(t, Y_{t+\varepsilon}^y\right)\right] - E[g(t, Y_t^y)]}{\varepsilon} = \end{aligned}$$

$$= E\left[g_t\left(t, Y_t^y\right)\right] + \lim_{\varepsilon \to 0} \frac{E\left[E\left[g\left(t, Y_{t+\varepsilon}^y\right)/\mathcal{F}_t\right]\right] - E\left[g\left(t, Y_t^y\right)\right]}{\varepsilon} =$$
$$= E\left[g_t\left(t, Y_t^y\right)\right] + \lim_{\varepsilon \to 0} \frac{E\left[E\left[g\left(t, Y_\varepsilon^\alpha\right)\right]\right|\alpha = Y_t^y\right] - E\left[g\left(t, Y_t^y\right)\right]}{\varepsilon} =$$
$$= E\left[g_t\left(t, Y_t^y\right)\right] + E\left[\lim_{\varepsilon \to 0} \frac{E\left[g\left(t, Y_\varepsilon^\alpha\right)\right] - g\left(t, \alpha\right)}{\varepsilon}\right| \alpha = Y_t^y\right] =$$
$$= E\left[g_t\left(t, \alpha\right) + Ag\left(t, \alpha\right)\right| \alpha = Y_t^y\right] = 0.$$

Since the derivative of $u\left(t, y\right)$ is zero, then $u\left(t, y\right)$ is independent of t and therefore $u\left(t, y\right) = u\left(0, y\right)$. So

$$E\left[g\left(t, Y_t^y\right)\right] = g\left(0, y\right).$$

We must show only, that point 1 follows from point 2. We have

$$E^Y\left[g\left(t+s, Y_{t+s}^y\right)\right|\mathcal{F}_s\right] = E^Y\left[g\left(t+s, Y_t^\alpha\right)\right]\Big|\,\alpha = Y_s^y.$$

Take the $E^Y\left[g\left(t+s, Y_t^\alpha\right)\right]$ in the Taylor's series

$$E^Y\left[g\left(t+s, Y_t^\alpha\right)\right] = E^Y\left[\sum_{n=0}^{\infty} \frac{s^n}{n!} g^{(n)}\left(t, Y_t^\alpha\right)\right] =$$
$$= \sum_{n=0}^{\infty} \frac{s^n}{n!} E^Y\left[g^{(n)}\left(t, Y_t^\alpha\right)\right].$$

So we must consider the expectations of the derivatives of g. We have

$$E\left[g_t\left(t, Y_t^y\right)\right] = E\left[\lim_{\varepsilon \to 0} \frac{g\left(t, Y_t^y\right) - g\left(t-\varepsilon, Y_t^y\right)}{\varepsilon}\right] =$$
$$= \lim_{\varepsilon \to 0} \frac{E\left[g\left(t, Y_t^y\right)\right] - E\left[g\left(t-\varepsilon, Y_t^y\right)\right]}{\varepsilon} =$$
$$= \lim_{\varepsilon \to 0} \frac{E\left[g\left(t, Y_t^y\right)\right] - E\left[E\left[g\left(t-\varepsilon, Y_t^y\right)\right]/\mathcal{F}_\varepsilon\right]}{\varepsilon} =$$
$$= \lim_{\varepsilon \to 0} \frac{E\left[g\left(t, Y_t^y\right)\right] - E\left[E\left[g\left(t-\varepsilon, Y_{t-\varepsilon}^\alpha\right)\right]\right|\alpha = Y_\varepsilon^y\right]}{\varepsilon} =$$
$$= \lim_{\varepsilon \to 0} \frac{g\left(0, y\right) - E\left[g\left(0, \alpha\right)\right|\alpha = Y_\varepsilon^y\right]}{\varepsilon}$$

i.e. $E^Y\left[g_t\left(t, Y_t^y\right)\right]$ is independent of t, and therefore $E^Y\left[g_t\left(t, Y_t^y\right)\right] = g_t\left(0, y\right)$. In fact we proved that if $E^Y\left[g\left(t, Y_t^y\right)\right]$ is independent from t, then $E^Y\left[g_t\left(t, Y_t^y\right)\right]$ is also independent of t. Analogously we can prove that $E^Y\left[g_t^{(n)}\left(t, Y_t^y\right)\right]$ is independent of t and therefore $E^Y\left[g_t^{(n)}\left(t, Y_t^y\right)\right] = g_t^{(n)}\left(0, y\right)$. Since

$$E^Y \left[g\left(t+s, Y_t^\alpha\right)\right] = \sum_{n=0}^{\infty} \frac{s^n}{n!} E^Y \left[g^{(n)}\left(t, Y_t^\alpha\right)\right],$$

then

$$\begin{aligned} E^Y \left[g\left(t+s, Y_t^\alpha\right)\right] &= \sum_{n=0}^{\infty} \frac{s^n}{n!} E^Y \left[g^{(n)}\left(t, Y_t^\alpha\right)\right] = \\ &= \sum_{n=0}^{\infty} \frac{s^n}{n!} g^{(n)}\left(0, \alpha\right) = \\ &= g\left(s, \alpha\right) \end{aligned}$$

from where we derive for the conditional expectation

$$\begin{aligned} E^Y \left[g\left(t+s, Y_{t+s}^y\right)\middle| \mathcal{F}_s\right] &= E^Y \left[g\left(t+s, Y_t^\alpha\right)\right]\middle| \alpha = Y_s^y = \\ &= g\left(s, \alpha\right)\middle| \alpha = Y_s^y = \\ &= g\left(s, Y_s^y\right) \end{aligned}$$

□

Note that the proofs of the propositions 8.2 and 8.3 are closely related because of the Markov property of the two types of processes.

8.3 Martingality and the Ito-Levy processes

Since the Ito-Levy processes are not markovian, the analogous proposition is not true. Let $Z_t^z = X_t^x\middle| x = Y_t^z$. Now we can see easily that if $g\left(Z_t^z\right)$ is a martingale, then

$$\begin{aligned} E\left[g\left(Z_t^z\right)\right] &= E\left[g\left(Z_t^z\right)/\mathcal{F}_0^Z\right] = \\ &= g\left(z\right). \end{aligned}$$

Therefore

$$A_t g\left(z\right) = \lim_{\varepsilon \to 0} \frac{E\left[g\left(Z_{t+\varepsilon}^z\right)\right] - E\left[g\left(Z_t^z\right)\right]}{\varepsilon} = 0.$$

So we see that if $g\left(Z_t^z\right)$ is a martingale, then the expectation is independent of the time and the infinitesimal generator of g is zero. With the next example we shall show, that it is possible for the infinitesimal generator to be zero, but $g\left(Z_t^z\right)$ is not a martingale.

Example 8.1. Let X_t^x be the solution of the stochastic differential equation

$$\begin{aligned} dX_t^x &= -bdt + \sigma X_t^x dB_t \\ X_0^x &= x. \end{aligned}$$

We can verify that

$$X_t^x = \exp\left(\sigma B_t - \frac{1}{2}\sigma^2 t\right)\left[x - b\int_0^t \exp\left(\sigma B_s - \frac{1}{2}\sigma^2 s\right) ds\right].$$

Since the expectation of the stochastic integral is zero, then $EX_t^x = x - bt$.

Let Y be a Levy-process with the generating triple $(\mu, 0, \Pi)$ and we want the measure $\Pi(\cdot)$ to be such that $E(Y) < \infty$. We shall set more conditions for μ later. We have for the infinitesimal generator of the function $g(z) = z$

$$\begin{aligned} A^Y g(z) &= \mu\frac{\partial}{\partial z}g(z) + \int_{-\infty}^{+\infty} g(z+y) - g(z) - I_D(y)y\frac{\partial}{\partial z}g(z)\Pi(dy) = \\ &= \mu + \int_{-\infty}^{+\infty} y - I_D(y)y\Pi(dy). \end{aligned}$$

We want μ to be such that $b = \mu + \int_{-\infty}^{+\infty} y - I_D(y)y\Pi(dy)$. We see that $A_0 g(z) = 0$. We have also

$$\begin{aligned} A_s g(z) &= E^Y\left[E^X\left\{A^X g(\alpha)\middle|\,\alpha = X_s^x\right\}\middle|\,x = Y_s^z\right] + \\ &+ E^Y\left\{A^Y E^X\left[g(X_s^\alpha)\right]\middle|\,\alpha = Y_s^z\right\} = \\ &= E^Y\left[E^X\left\{-b\middle|\,\alpha = X_s^x\right\}\middle|\,x = Y_s^z\right] + E^Y\left\{A^Y(\alpha - bs)\middle|\,\alpha = Y_s^z\right\} = \\ &= -b + E^Y\left\{\mu + \int_{-\infty}^{+\infty} y - I_D(y)y\Pi(dy)\middle|\,\alpha = Y_s^z\right\} = \\ &= 0. \end{aligned}$$

We shall show that the process $Z_t^z = X_t^x|\, x = Y_t^z$ is not a martingale. We have for the conditional expectation

$$\begin{aligned} &E\left[Z_{t+s}^z/\mathcal{F}_s\right]\left(\omega^X, \omega^Y\right) = \\ &\quad = E^Y\left[\left.E^X\left[X_t^\alpha\right]\middle|\,\alpha = X_s^x(\omega^X)\right|x = Y_t^\beta\right]\Big|\,\beta = Y_s^z(\omega^Y) = \\ &\quad = E^Y\left[\left.\alpha - bt\middle|\,\alpha = X_s^x(\omega^X)\right|x = Y_t^\beta\right]\Big|\,\beta = Y_s^z(\omega^Y) = \\ &\quad = E^Y\left[X_s^x(\omega^X) - bt\middle|\,x = Y_t^\beta\right]\Big|\,\beta = Y_s^z(\omega^Y). \end{aligned}$$

Suppose that Z is a martingale. Then

$$E^Y\left[X_s^x(\omega^X)-bt\big|\,x=Y_t^\beta\right]=X_s^x(\omega^X)$$

for arbitrary s,t and ω^X. So we see that for arbitrary s and ω^X

$$E^Y\left[X_s^x(\omega^X)-bt\big|\,x=Y_t^\beta\right]$$

is independent of t. Using proposition 8.3 we conclude that $X_s^x(\omega^X)-bt\big|\,x=Y_t$ is a Y-martingale for arbitrary s and ω^X. We have from the form of $X_s^x(\omega^X)$ that

$$\begin{aligned}
&-bt+X_s^x(\omega^X)\big|\,x=Y_t=\\
&\quad=\exp\left(\sigma B_s-\frac{1}{2}\sigma^2 s\right)\left[Y_t-b\int\limits_0^s\exp\left(\sigma B_u-\frac{1}{2}\sigma^2 u\right)du\right]-bt=\\
&\quad=\exp\left(\sigma B_s-\frac{1}{2}\sigma^2 s\right)Y_t-\\
&\quad-b\exp\left(\sigma B_s-\frac{1}{2}\sigma^2 s\right)\int\limits_0^s\exp\left(\sigma B_u-\frac{1}{2}\sigma^2 u\right)du-bt
\end{aligned}$$

is a Y-martingale for arbitrary s and ω^X . Since

$$\exp\left(\sigma B_s-\frac{1}{2}\sigma^2 s\right)$$

and

$$b\exp\left(\sigma B_s-\frac{1}{2}\sigma^2 s\right)\int\limits_0^s\exp\left(\sigma B_u-\frac{1}{2}\sigma^2 u\right)du$$

may have arbitrary nonnegative values, then $\alpha Y_t-\beta-bt$ is a martingale for arbitrary nonnegative α and β . Using proposition 8.3(points 1 and 3) we conclude that

$$0=\alpha\left[\mu+\int\limits_{-\infty}^{+\infty}y-I_D(y)y\Pi(dy)\right]-b=b\,(\alpha-1)$$

which is not true. This proves that Z is not a martingale.

References

[1] J. Bertoin, *Levy Processes*, Cambridge University Press, 1976

[2] S. I. Boyarchenko and S. Z. Levendorskii, *Non-Gaussian Merton-Black-Scholes Theory*,World Scientific, New Jersey, London, Singapore, Hong Kong, 2002

[3] R. Cont, P. Tankov, *Financial modeling with jump processes*, Chapman & Hall 2004

[4] J. Lamperti, *Stochastic processes*, Springer-Verlag 1977

[5] L. Nirenberg, *Lectures on linear partial differential equations*, American Mathematical Society, 1973

[6] B. Oksendal,*Stochastic differential equations: an introduction with aplication*, Springer-Verlag 1998

[7] S. Rachev, S. Mittnik,*Stable Paretian Models in Finance*, John Wiley&Sons Ltd, 2000

[8] G. Samorodnitsky and M. S. Taqqu, *Stable Non-Gaussian Random Processes*, Chapman & Hall, New York, London

[9] K. Sato, *Levy processes and infinitely divisible distributions*, Cambridge University Press 1999

JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS,VOL.5,NO.4,299-322,2007,COPYRIGHT 2007 EUDOXUS PRESS,LLC

Changing the Probability Measure for the Ito-Levy Processes

Tsvetelin Stefanov Zaevski
Department of Mathematics and Informatics
Sofia University "St. Kliment Ohridski"
Sofia, Bulgaria
e-mail: t_s_zaevski@abv.bg

November 15, 2006

Abstract

In this article we study the problem of changing the probability measure when the processes are of the Ito-Levy type (defined in [10]).In view of the nature of these processes we must generalize the concept of equivalent measures. We shall get the results for the expectation, the conditional expectation and for the infinitisimal genrator w.r.t. the new measure.

As a corollary of the received method for changing the measures, we build an algorithm for constructing equivalent martingale measures and show that it is the unique one of this type.

In next papers we apply these processes to financial modeling, so the question of changing the probability measure is basic for the results proved there.
Keywords: Ito-Levy process, probability measure, Radon-Nikodim derivative, agreement condition, martingality.

1 Equivalent Probability Measure. Generalization.

Let X_t^x be an Ito process, which is a solution of the stochastic differential equation:

$$\begin{aligned} dX_t^x &= \mu\left(X_t^x\right)dt + \sigma\left(X_t^x\right)dB_t \\ X_0^x &= x, x \in \mathbf{R}, \end{aligned} \tag{1}$$

where B_t is a standard Brownian motion. We have from [6, theorem 5.2.1.] sufficient conditions for the existence and uniqueness of the solution.

Also let Y_t be a Levi process, independent of X_t, with a generating triple $(b, 0, \Pi)$. We denote the probability spaces of the processes X_t and Y_t by

$\left(\Omega^X, \mathcal{F}^X, P^X\right)$ and $\left(\Omega^Y, \mathcal{F}^Y, P^Y\right)$ and the corresponding expectations by E^X and E^Y. Since the Levy processes are heavy tail processes, then when we examine the expectation $E\left(g\left(Y_t\right)\right)$ we assume that $g\left(\cdot\right)$ is a suitable function.

Let Z be the Ito-Levy process

$$Z_t^z(\omega) = Z_t^z(\omega^X, \omega^Y) = X_t^x(\omega^X)\big|\, x = Y_t^z(\omega^Y),$$

defined in [10]. We denote its probability space by $\left(\Omega^Z, \mathcal{F}^Z, P^Z\right)$.

We shall study the question of changing the probability measures. We shall generalize the concept of equivalent measures and show how change the expectations and the infinitesimal generator under the new measure.

Let us transform the expectation (see [10, definition 1.1] for $P\left(Z_t^z < b\right)$) :

$$\begin{aligned}
E\left(h(Z_t^z)\right) &= \int_{-\infty}^{+\infty} h(b) dP\left(Z_t^z < b\right) = \\
&= \int_{-\infty}^{+\infty} h(b) dE^Y\left[P\left(X_t^x < b\right)\big|\, x = Y_t^z\right] = \\
&= \int_{-\infty}^{+\infty} h(b) d\int_{-\infty}^{+\infty} P^X\left(X_t^x < b\right) dP^Y\left(Y_t^z < x\right) = \\
&= \int_{-\infty}^{+\infty} h(b) d\int_{-\infty}^{+\infty} \int_{\omega^X: X_t^x(\omega^X) < b} 1 P^X\left(d\omega^X\right) dP^Y\left(Y_t^z < x\right) = \\
&= \int_{-\infty}^{+\infty} h(b) \int_{-\infty}^{+\infty} \int_{\omega^X: X_t^x(\omega^X) = b} P^Y\left(Y_t^z = x\right) P^X\left(d\omega^X\right) dx db. \qquad (2)
\end{aligned}$$

Let us notice that we use

$$\int_{\omega^X: X_t^x(\omega^X) = b} (\circ)\, P^X\left(d\omega^X\right)$$

instead of

$$\frac{\partial}{\partial b} \int_{\omega^X: X_t^x(\omega^X) < b} (\circ)\, P^X\left(d\omega^X\right).$$

and $P^Y\left(Y_t^z = x\right)$ for the density of the process Y_t^z.

If we change the measures P^X with Q^X and P^Y with Q^Y in the formula (2), then there will be no principle change. Let us discuss the change of P^X with Q^X and a set of equivalent to P^Y measures, ${}^{\omega^X}Q^Y$. If for identical w.r.t.

the σ-algebra $\mathcal{F}_t^X$ states ω^X, the measures ${}^{\omega^X}Q^Y$ coincide, then we shall say that ${}^{\omega^X}Q^Y$ is $\mathcal{F}_t^X$- measurable. In other words ${}^{\omega^X}Q^Y$ is $\mathcal{F}_t^X$- measurable if $\xi\left(\omega^X\right) = {}^{\omega^X}E^Y[\eta]$ is $\mathcal{F}_t^X$- measurable for every $\mathcal{F}^Y$ -measurable η. Let ${}^{\omega^X}Q^Y$ be $\mathcal{F}_t^X$-measurable.

Using $\omega^X \to {}^{\omega^X}Q^Y$ we shall define the new measure Q, based on Q^X and ${}^{\omega^X}Q^Y$. Thus the expectation will be

$$E\left(h(Z_t^z)\right) = \int_{-\infty}^{+\infty} h(b) \int_{-\infty}^{+\infty} \int_{\omega^X : X_t^x(\omega^X)=b} {}^{\omega^X}Q^Y\left(Y_t^z = x\right) Q^X\left(d\omega^X\right) dx db.$$

In view of the previous notation this is

$$E\left(h(Z_t^z)\right) = \int_{-\infty}^{+\infty} h(b) \int_{-\infty}^{+\infty} \frac{\partial}{\partial b} \int_{\omega^X : X_t^x(\omega^X)<b} {}^{\omega^X}Q^Y\left(Y_t^z = x\right) Q^X\left(d\omega^X\right) dx db.$$

We see that if the set of measures ${}^{\omega^X}Q^Y$ is fixed, then there is no change in the form of the expectation.

For arbitrary $s_1 < s_2 < ... < s_l$ and Borel sets $B_1, B_2, ..., B_l$ we define the probability measure

$$Q\left(Z_{s_1}^z \in B_1, Z_{s_2}^z \in B_2, ..., Z_{s_l}^z \in B_l\right) =$$
$$= \int_{x \in R^l} \int_{\omega^X : \left\{\begin{array}{l} X_{s_1}^{x_1}\left(\omega^X\right) \in B_1, \\ X_{s_2}^{x_2}\left(\omega^X\right) \in B_2, \\ ,..., \\ X_{s_l}^{x_l}\left(\omega^X\right) \in B_l \end{array}\right\}} {}^{\omega^X}Q^Y \left(\begin{array}{l} Y_{s_1}^z = x_1, \\ Y_{s_2}^z = x_2, \\ ,..., \\ Y_{s_l}^z = x_l \end{array}\right) Q^X\left(d\omega^X\right) dx_1 ... dx_l. \tag{3}$$

We shall prove that Q really defines a probability measure:

1. $Q\left(Z_{s_1}^z \in B_1, Z_{s_2}^z \in B_2, ..., Z_{s_l}^z \in B_l\right) \geq 0$ by definition (3).
2. To verify that $Q(\Omega) = 1$ it is sufficient to show that $Q\left(Z_s^z \in R\right) = 1$:

$$Q\left(Z_s^z \in R\right) = \int_{-\infty}^{+\infty} \int_{\omega^X : X_s^x \in R} {}^{\omega^X}Q^Y\left(Y_s^z = x\right) Q^X\left(d\omega^X\right) dx =$$
$$= \int_{-\infty}^{+\infty} \int_{\omega^X \in \Omega^X} {}^{\omega^X}Q^Y\left(Y_s^z = x\right) Q^X\left(d\omega^X\right) dx =$$

$$= \int_{-\infty}^{+\infty} E^X \left[{}^{\omega^X} Q^Y \left(Y_s^z = x \right) \right] dx =$$
$$= E^X \left[\int_{-\infty}^{+\infty} {}^{\omega^X} Q^Y \left(Y_s^z = x \right) dx \right] =$$
$$= E^X [1] = 1.$$

3. Let us consider two disjoint sets of the σ-algebra $\mathcal{F}$, i.e. the corresponding Borel sets are disjoint for some moment from $s_1, s_2, ..., s_l$. Also the coresponding α's is identical for the two sets. Without any restriction we may think that this is s_l. So the two elements will be of the type

$$\left\{ Z_{s_1}^{\alpha_1} \in B_1, Z_{s_2}^{\alpha_2} \in B_2, ..., Z_{s_l}^{\alpha_l} \in B_l \right\},$$
$$\left\{ Z_{s_1}^{\overline{\alpha_1}} \in \overline{B_1}, Z_{s_2}^{\overline{\alpha_2}} \in \overline{B_2}, ..., Z_{s_l}^{\alpha_l} \in \overline{B_l} \right\}.$$

We can easily verify that

$$Q \left(Z_{s_1}^z \in B_1, Z_{s_2}^z \in B_2, ..., Z_{s_l}^z \in B_l \right) +$$
$$+ Q \left(Z_{s_1}^z \in \overline{B_1}, Z_{s_2}^z \in \overline{B_2}, ..., Z_{s_l}^z \in \overline{B_l} \right) =$$
$$= \int_{x \in \mathbf{R}^l} \int_{\omega^X : \left\{ \begin{array}{l} X_{s_1}^{x_1} \left(\omega^X \right) \in B_1 \cup \overline{B_1}, \\ X_{s_2}^{x_2} \left(\omega^X \right) \in B_2 \cup \overline{B_2}, \\ , ..., \\ X_{s_l}^{x_l} \left(\omega^X \right) \in B_l \cup \overline{B_l} \end{array} \right\}} {}^{\omega^X} Q^Y \begin{pmatrix} Y_{s_1}^z = x_1, \\ Y_{s_2}^z = x_2, \\ , ..., \\ Y_{s_l}^z = x_l \end{pmatrix} Q^X \left(d\omega^X \right) dx =$$
$$= Q \left(Z_{s_1}^z \in B_1 \cup \overline{B_1}, Z_{s_2}^z \in B_2 \cup \overline{B_2}, ..., Z_{s_l}^z \in B_l \cup \overline{B_l} \right).$$

4. The continuity of the probability measure comes from the possibility to interchange the sign of limit and Q in $\lim_{i \to \infty} Q(A_i)$ (A_i are inner sets of the σ- algebra $\mathcal{F}$, going to the empty set).

Proposition 1.1. *The measure Q, defined above, is equivalent to the natural measure P.*

Proof. Let for some A from the σ-algebra $\mathcal{F}$ we have that $P(A) = 0$. Since P is equivalent to ${}^{\omega^X}Q$ for every ω^X, then ${}^{\omega^X}Q(A) = 0$. Let B be the corresponding Borel set for Z_s^z. Therefore

$$0 = {}^{\omega^X}Q\left(Z^z_{s_1} \in B_1, Z^z_{s_2} \in B_2, ..., Z^z_{s_l} \in B_l\right) =$$
$$= {}^{\omega^X}E^Y\left[\begin{array}{l} Q^X\left(X^{x_1}_{s_1} \in B_1, X^{x_2}_{s_2} \in B_2, ..., X^{x_l}_{s_l} \in B_l\right)\big| \\ \big|x_1 = Y^z_{s_1}, x_2 = Y^z_{s_2}, ..., x_l = Y^z_{s_l} \end{array}\right]. \tag{4}$$

We shall examine two cases: when the process Y has a discrete distribution or when it has a continuous distribution.

We shall say that a point $x \in R^l$ is a mass point if the process Y^z_s (s is the set of moments $s_1, s_2, ..., s_l$) has a mass in it. Since the measures ${}^{\omega^X}Q^Y$ are equivalent to P^Y, then if one point is a mass point w.r.t. the measure P^Y, then it is a mass point w.r.t all measures ${}^{\omega^X}Q^Y$. From (4) we have that if x is a mass point, then

$$Q^X\left(X^{x_1}_{s_1} \in B_1, X^{x_2}_{s_2} \in B_2, ..., X^{x_l}_{s_l} \in B_l\right) = 0.$$

Let us look at the formula (3) for the measure Q. We shall show that in this case the inner integral

$$\int\limits_{\omega^X : \left\{\begin{array}{l} X^{x_1}_{s_1}\left(\omega^X\right) \in B_1, \\ X^{x_2}_{s_2}\left(\omega^X\right) \in B_2, \\ , ..., \\ X^{x_l}_{s_l}\left(\omega^X\right) \in B_l \end{array}\right\}} {}^{\omega^X}Q^Y\left(Y^z_{s_1} = x_1, Y^z_{s_2} = x_2, ...Y^z_{s_l} = x_l\right) Q^X\left(d\omega^X\right)$$

is equal to zero. If x is not a mass point, then

$${}^{\omega^X}Q^Y\left(Y^z_{s_1} = x_1, Y^z_{s_2} = x_2, ..., Y^z_{s_l} = x_l\right) = 0$$

for all ω^X. On the other hand if x is a mass point, then

$$Q^X\left(\omega^X : X^{x_1}_{s_1}\left(\omega^X\right) \in B_1, X^{x_2}_{s_2}\left(\omega^X\right) \in B_2, ..., X^{x_l}_{s_l}\left(\omega^X\right) \in B_l\right) = 0.$$

Therefore the integral is zero. So $Q\left(Z^z_{s_1} \in B_1, Z^z_{s_2} \in B_2, ..., Z^z_{s_l} \in B_l\right) = 0$.

If the process Y has a continuous distribution, then we have from (4) that for every ω^X

$$0 = {}^{\omega^X}E^Y\left[\begin{array}{l} Q^X\left(X^{x_1}_{s_1} \in B_1, X^{x_2}_{s_2} \in B_2, ..., X^{x_l}_{s_l} \in B_l\right)\big| \\ \big|x_1 = Y^z_{s_1}, x_2 = Y^z_{s_2}, ..., x_l = Y^z_{s_l} \end{array}\right] =$$
$$= \int\limits_{R_l} Q^X\left(\begin{array}{l} X^{x_1}_{s_1} \in B_1, \\ X^{x_2}_{s_2} \in B_2, ..., X^{x_l}_{s_l} \in B_l \end{array}\right) {}^{\omega^X}Q^Y\left(\begin{array}{l} Y^z_{s_1} = x_1, \\ Y^z_{s_2} = x_2, ..., Y^z_{s_l} = x_l \end{array}\right) dx.$$

Since Y has a continuous distribution, then

$$Q^X\left(X_{s_1}^{x_1}\in B_1,X_{s_2}^{x_2}\in B_2,...,X_{s_l}^{x_l}\in B_l\right)=0.$$

almost everywhere. In the same way as above we have that

$$Q\left(Z_{s_1}^{z}\in B_1,Z_{s_2}^{z}\in B_2,...,Z_{s_l}^{z}\in B_l\right)=0.$$

Reverse: let for some element A from the σ-algebra $\mathcal{F}$ we have $Q(A)=0$. We shall use the definition of the measure Q -(3). We shall study again separately the two cases for the distribution of the process Y. Suppose that Y has a discrete distribution. Since the probability (3) is equal to zero, then for all mass points x we have

$$Q^X\left(\omega^X:X_{s_1}^{x_1}\left(\omega^X\right)\in B_1,X_{s_2}^{x_2}\left(\omega^X\right)\in B_2,...,X_{s_l}^{x_l}\left(\omega^X\right)\in B_l\right)=0.$$

So for an arbitrary ω^X we have

$$^{\omega^X}E^Y\left[\begin{array}{l}Q^X\left(X_{s_1}^{x_1}\in B_1,X_{s_2}^{x_2}\in B_2,...,X_{s_l}^{x_l}\in B_l\right)\big|\\ \big|x_1=Y_{s_1}^{z},x_2=Y_{s_2}^{z},...,x_l=Y_{s_l}^{z}\end{array}\right]=0,$$

i.e. $^{\omega^X}Q\left(Z_{s_1}^{z}\in B_1,Z_{s_2}^{z}\in B_2,...,Z_{s_l}^{z}\in B_l\right)=0$. Since the measures $^{\omega^X}Q$ and P are equivalent, then we have that $P(A)=0$.

We must only examine the case when the process Y has a continuous distribution. Then for all measures $^{\omega^X}Q^Y$ we have

$$^{\omega^X}Q^Y\left(Y_{s_1}^{z}=x_1,Y_{s_2}^{z}=x_2,...,Y_{s_l}^{z}=x_l\right)>0.$$

Since

$$Q\left(Z_{s_1}^{z}\in B_1,Z_{s_2}^{z}\in B_2,...,Z_{s_l}^{z}\in B_l\right)=$$
$$=\int\limits_{x\in R^l}\;\int\limits_{\omega^X:\left\{\begin{array}{l}X_{s_1}^{x_1}\left(\omega^X\right)\in B_1,\\ ,X_{s_2}^{x_2}\left(\omega^X\right)\in B_2,\\ ,...,\\ ,X_{s_l}^{x_l}\left(\omega^X\right)\in B_l\end{array}\right\}}{}^{\omega^X}Q^Y\left(\begin{array}{l}Y_{s_1}^{z}=x_1,\\ ,Y_{s_2}^{z}=x_2,\\ ,...,\\ ,Y_{s_l}^{z}=x_l\end{array}\right)Q^X\left(d\omega^X\right)dx_1...dx_l=0,$$

then

$$Q^X\left(\omega^X:X_{s_1}^{x_1}\left(\omega^X\right)\in B_1,X_{s_2}^{x_2}\left(\omega^X\right)\in B_2,...,X_{s_l}^{x_l}\left(\omega^X\right)\in B_l\right)=0$$

almost everywhere and therefore

$$
{}^{\omega^X}Q(A) = {}^{\omega^X}E^Y \left[\begin{array}{l} Q^X \left(X_{s_1}^{x_1} \in B_1, X_{s_2}^{x_2} \in B_2, ..., X_{s_l}^{x_l} \in B_l \right) \big| \\ \big| x_1 = Y_{s_1}^z, x_2 = Y_{s_2}^z, ..., x_l = Y_{s_l}^z \end{array} \right] =
$$

$$
= \int_{R_l} Q^X \begin{pmatrix} X_{s_1}^{x_1} \in B_1, \\ X_{s_2}^{x_2} \in B_2, \\ , ..., X_{s_l}^{x_l} \in B_l \end{pmatrix} {}^{\omega^X}Q^Y \begin{pmatrix} Y_{s_1}^z = x_1, \\ Y_{s_2}^z = x_2, \\ , ..., Y_{s_l}^z = x_l \end{pmatrix} dx = 0.
$$

We see that again $P(A) = 0$.

We can conclude now that P and Q are equivalent measures.

$\square$

Analogously we can define the probability measure for the σ-algebra generated from several different Ito-Levy processes.

Let examine the next example.

Example 1.1.

$$
Q(X_t^z \in A, Y_t \in B) = Q(X_t^{x_1} | \, x_1 = z \in A, x_2 | \, x_2 = Y_t \in B) =
$$

$$
= \int_{x \in R^2} \int_{\omega^x \in \left\{ \begin{smallmatrix} X_t^{x_1} \in A, \\ x_2 \in B \end{smallmatrix} \right\}} {}^{\omega^X}Q^Y \begin{pmatrix} z = x_1, \\ Y_t = x_2 \end{pmatrix} Q^X \left(d\omega^X \right) dx =
$$

$$
= \int_{\omega^x \in \{X_t^z \in A\}} {}^{\omega^X}Q^Y (Y_t \in B) Q^X \left(d\omega^X \right)
$$

Using this example we can define the probability measure for the σ-algebra $\mathcal{H}$, generated from the processes X_t^x for all x and Y_t. Also we conclude that

$$
dQ(\omega) = d^{\omega^X}Q^Y \left(\omega^Y \right) dQ \left(\omega^X \right)
$$

Now we shall study the question of the expectation, the conditional expectation and the infinitesimal generator w.r.t. the measure, so defined.

1.1 Expectation

Proposition 1.2.

$$
E(h(Z_s^z)) = \int_{-\infty}^{+\infty} h(b) \int_{-\infty}^{+\infty} \int_{\omega^X : X_s^x(\omega^X) = b} {}^{\omega^X}Q^Y (Y_s^z = x) \, Q^X \left(d\omega^X \right) dx db.
$$

Proof.

$$
E(h(Z_s^z)) = \int_{-\infty}^{+\infty} h(b) dQ(Z_s^z < b) =
$$

$$= \int_{-\infty}^{+\infty} h(b) d \int_{-\infty}^{+\infty} \int_{\omega^X : X_s^x(\omega^X) < b} {}^{\omega^X} Q^Y \left(Y_s^z = x \right) Q^X \left(d\omega^X \right) dx =$$

$$= \int_{-\infty}^{+\infty} h(b) \int_{-\infty}^{+\infty} \frac{\partial}{\partial b} \int_{\omega^X : X_s^x(\omega^X) < b} {}^{\omega^X} Q^Y \left(Y_s^z = x \right) Q^X \left(d\omega^X \right) dx db =$$

$$= \int_{-\infty}^{+\infty} h(b) \int_{-\infty}^{+\infty} \int_{\omega^X : X_s^x(\omega^X) = b} {}^{\omega^X} Q^Y \left(Y_s^z = x \right) Q^X \left(d\omega^X \right) dx db$$

as we expected from the beginning.

□

Proposition 1.3.

$$E \left(h(Z_s^z) \right) = E^X \left[{}^{\omega^X} E^Y \left[h(X_s^x) | \, x = Y_s^z \right] \right].$$

Proof.

$$E \left(h(Z_s^z) \right) = \int_{-\infty}^{+\infty} h(b) \int_{-\infty}^{+\infty} \int_{\omega^X : X_s^x(\omega^X) = b} {}^{\omega^X} Q^Y \left(Y_s^z = x \right) Q^X \left(d\omega^X \right) dx db =$$

$$= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{\omega^X : X_s^x(\omega^X) = b} h(b)^{\omega^X} Q^Y \left(Y_s^z = x \right) Q^X \left(d\omega^X \right) db dx =$$

$$= \int_{-\infty}^{+\infty} \int_{\omega^X \in \Omega^X} h(X_s^x)^{\omega^X} Q^Y \left(Y_s^z = x \right) Q^X \left(d\omega^X \right) dx =$$

$$= \int_{-\infty}^{+\infty} E^X \left[h(X_s^x)^{\omega^X} Q^Y \left(Y_s^z = x \right) \right] dx =$$

$$= E^X \left[\int_{-\infty}^{+\infty} h(X_s^x)^{\omega^X} Q^Y \left(Y_s^z = x \right) dx \right] =$$

$$= E^X \left[{}^{\omega^X} E^Y \left[h(X_s^x) | \, x = Y_s^z \right] \right].$$

□

1.2 Conditional Expectation

Proposition 1.4. *Let $t \le s \le T$. Then*

$$\int\limits_{\omega \in B} g\left(Z_T^z\right) Q\left(d\omega\right) =$$
$$= \int\limits_{\omega \in B} {}^{\omega^X} E^Y \left[E^X \left[g\left(X_{T-s}^{\alpha}\right) \right] \middle| \alpha = X_s^x(\omega^X) \middle| x = Y_{T-s}^{\beta} \right] \middle| \beta = Y_s^z(\omega^Y) Q\left(d\omega\right)$$

for every set $B \in \mathcal{F}_s$.

Proof. Let $B \in \mathcal{F}_s$. Define

$$B^X\left(x\right) = \left(X_s^x\right)^{-1}\left(Z_s^z(B)\right)$$
$$B^Y\left(x\right) = \left\{\omega^Y : Y_s^z = x\right\}.$$

We note that $B^X\left(x\right) \in \mathcal{F}_s^X$ and $B^Y\left(x\right) \in \mathcal{F}_s^Y$. Let us consider:

$$\int\limits_{\omega \in B} g\left(Z_T^z(\omega)\right) Q(d\omega) =$$
$$= \int\limits_{-\infty}^{+\infty} \int\limits_{\omega^X \in B^X(x)} \left[\int\limits_{\omega^Y \in B^Y(x)} g\left(X_T^{\gamma}(\omega^X)\right) \middle| \gamma = Y_T^z(\omega^Y)^{\omega^X} Q^Y(d\omega^Y) \right] Q^X(d\omega^X) dx.$$

Let us choose $\overline{Q}^Y$, one of the measures ${}^{\omega^X}Q^Y$, for index and then the corresponding Radon-Nikodim derivatives are $\xi\left(\omega^X\right) = \frac{d^{\omega^X} Q^Y\left(d\omega^Y\right)}{d\overline{Q}^Y(d\omega^Y)}$. Denote that $\xi\left(\omega^X\right)$ is dependent on ω^Y, but this dependence is not essential for the proof. So

$$\int\limits_{\omega \in B} g\left(Z_T^z(\omega)\right) Q(d\omega) =$$

$$= \int\limits_{-\infty}^{+\infty} \int\limits_{\omega^X \in B^X(x)} \int\limits_{\omega^Y \in B^Y(x)} g\left(X_T^{\gamma}(\omega^X)\right) \middle| \gamma = Y_T^z(\omega^Y) \xi\left(\omega^X\right) \overline{Q}^Y(d\omega^Y) Q^X(d\omega^X) dx.$$

Changing the order of integration, using the Markov property of X and the fact that ${}^{\omega^X}Q^Y$ is $\mathcal{F}_t^X$-measurable measure and therefore $\xi\left(\omega^X\right)$ is $\mathcal{F}_t^X$-measurable, we obtain:

$$\int\limits_{\omega \in B} g\left(Z_T^z(\omega)\right) Q(d\omega) =$$

$$
= \int_{-\infty}^{+\infty} \int_{\omega^Y \in B^Y(x)} \left(\int_{\substack{\omega^X \in B^X(x) \\ |\gamma = Y_T^z(\omega^Y)}} \begin{pmatrix} \xi\left(\omega^X\right). \\ .g\left(X_T^\gamma(\omega^X)\right) \end{pmatrix} Q^X(d\omega^X) \Bigg| \right) \overline{Q}^Y(d\omega^Y)dx =
$$

$$
= \int_{-\infty}^{+\infty} \int_{\omega^Y \in B^Y(x)} \left[\int_{\substack{\omega^X \in B^X(x) \\ |\gamma = Y_T^z(\omega^Y)}} \left[\begin{pmatrix} \xi\left(\omega^X\right). \\ .E^X\left[g\left(X_{T-s}^\alpha\right)\right] \end{pmatrix} \Bigg| \; \Big| \alpha = X_s^\gamma(\omega^X) \right] Q^X(d\omega^X) \Bigg| \right] \overline{Q}^Y(d\omega^Y)dx.
$$

Changing the order of integration again, we set:

$$
\int_{\omega \in B} g\left(Z_T^z(\omega)\right) Q(d\omega) = \tag{5}
$$

$$
= \int_{-\infty}^{+\infty} \int_{\omega^X \in B^X(x)} \left[\int_{\omega^Y \in B^Y(x)} \left[\begin{pmatrix} \xi\left(\omega^X\right). \\ .E^X\left[g\left(X_{T-s}^\alpha\right)\right] | \\ |\alpha = X_s^\gamma(\omega^X) \end{pmatrix} \Bigg| \; \Big| \gamma = Y_T^z(\omega^Y) \right] \overline{Q}^Y(d\omega^Y) \right] Q^X(d\omega^X)dx =
$$

$$
= \int_{-\infty}^{+\infty} \int_{\omega^X \in B^X(x)} \left[\int_{\omega^Y \in B^Y(x)} \left[\begin{pmatrix} E^X\left[g\left(X_{T-s}^\alpha\right)\right] | \\ |\alpha = X_s^\gamma(\omega^X) \end{pmatrix} \Bigg| \; \Big| \gamma = Y_T^z(\omega^Y) \right]^{\omega^X} Q^Y(d\omega^Y) \right] Q^X(d\omega^X)dx. \tag{6}
$$

Using the Markov property of Y

$$
E^Y\left(h(Y_T^z)/\mathcal{F}_s^Y\right) = E^Y\left(h(Y_{T-s}^\alpha)\right)\big|\,\alpha = Y_s^z
$$

for

$$
h(x) = E^X\left[g\left(X_{T-s}^\alpha\right)\right]\big|\,\alpha = X_s^x(\omega^X)
$$

we obtain

$$
\int_{\omega^Y \in B^Y(x)} E^X\left[g\left(X_{T-s}^\alpha\right)\right]\big|\,\alpha = X_s^\gamma(\omega^X)\big|\,\gamma = Y_T^z(\omega^Y)^{\omega^X} Q^Y(d\omega^Y) =
$$

$$
= \int_{\omega^Y \in B^Y(x)} {}^{\omega^X} E^Y\left[\begin{pmatrix} E^X\left[g\left(X_{T-s}^\alpha\right)\right] | \\ |\alpha = X_s^\gamma(\omega^X) \end{pmatrix} \Bigg| \gamma = Y_{T-s}^\beta \right] \Bigg| \beta = Y_s^z(\omega^Y)^{\omega^X} Q^Y(d\omega^Y). \tag{7}
$$

Replacing (7) in (6) we obtain

$$\int\limits_{\omega\in B} g\left(Z_T^z\right)Q\left(d\omega\right) =$$

$$= \int\limits_{-\infty}^{+\infty}\int\limits_{\omega^X\in B^X(x)}\left[\int\limits_{\omega^Y\in B^Y(x)}\left[\begin{array}{l}\left(\begin{array}{l}E^X\left[g\left(X_{T-s}^{\alpha}\right)\right]\big| \\ \big|\alpha = X_s^{\gamma}(\omega^X)\end{array}\right)\Bigg| \\ \big|\gamma = Y_T^z(\omega^Y)\end{array}\right]{}^{\omega^X}Q^Y(d\omega^Y)\right]Q^X(d\omega^X)dx =$$

$$= \int\limits_{-\infty}^{+\infty}\int\limits_{\omega^X\in B^X(x)}\left[\int\limits_{\omega^Y\in B^Y(x)}\begin{array}{l}{}^{\omega^X}E^Y\left[\begin{array}{l}\left(\begin{array}{l}E^X\left[g\left(X_{T-s}^{\alpha}\right)\right]\big| \\ \big|\alpha = X_s^{\gamma}(\omega^X)\end{array}\right)\Bigg| \\ \Big|\gamma = Y_{T-s}^{\beta}\end{array}\right]\Bigg| \\ \big|\beta = Y_s^z(\omega^Y)\ {}^{\omega^X}Q^Y(d\omega^Y)\end{array}\right]Q^X(d\omega^X)dx =$$

$$= \int\limits_{\omega\in B}{}^{\omega^X}E^Y\left[E^X\left[g\left(X_{T-s}^{\alpha}\right)\right]\big|\,\alpha = X_s^{\gamma}(\omega^X)\big|\,\gamma = Y_{T-s}^{\beta}\right]\Big|\,\beta = Y_s^z(\omega^Y)Q\left(d\omega\right).$$

□

Corollary 1.1. *Let $t \le s \le T$. If*

$${}^{\omega^X}E^Y\left[E^X\left[g\left(X_{T-s}^{\alpha}\right)\right]\big|\,\alpha = X_s^{x}(\omega^X)\big|\,x = Y_{T-s}^{\beta}\right]\Big|\,\beta = Y_s^z(\omega^Y)$$

is $\mathcal{F}_s$ -measurable, then

$$E^Q\left[g\left(Z_T^z\right)\big|\,\mathcal{F}_s\right] = {}^{\omega^X}E^Y\left[E^X\left[g\left(X_{T-s}^{\alpha}\right)\right]\big|\,\alpha = X_s^{x}(\omega^X)\big|\,x = Y_{T-s}^{\beta}\right]\Big|\,\beta = Y_s^z(\omega^Y).$$

Corollary 1.2. *If the measures ${}^{\omega^X}Q^Y$ are such that for an arbitrary s*

$$E^X\left[g\left(X_{T-s}^{\alpha}\right)\right]\big|\,\alpha = X_s^{x}(\omega^X)\big|\,x = Y^{\beta}(\omega^Y)$$

are martingales w.r.t. them, then

$$E^Q\left[g\left(Z_T^z\right)\big|\,\mathcal{F}_s\right] = {}^{\omega^X}E^Y\left[E^X\left[g\left(X_{T-s}^{\alpha}\right)\right]\big|\,\alpha = X_s^{x}(\omega^X)\big|\,x = Y_{T-s}^{\beta}\right]\Big|\,\beta = Y_s^z(\omega^Y).$$

Proof. Since

$$E^X\left[g\left(X_{T-s}^{\alpha}(\omega^X)\right)\right]\big|\,\alpha = X_s^{x}(\omega^X)\big|\,x = Y^{\beta}(\omega^Y)$$

is a martingale, then

$${}^{\omega^X}E^Y\left[E^X\left[g\left(X_{T-s}^{\alpha}\right)\right]\big|\,\alpha = X_s^{x}(\omega^X)\big|\,x = Y_{T-s}^{\beta}\right]\Big|\,\beta = Y_s^z(\omega^Y) =$$
$$= E^X\left[g\left(X_{T-s}^{\alpha}\right)\right]\big|\,\alpha = X_s^{x}(\omega^X)\big|\,x = Y_s^z(\omega^Y)$$

and therefore is $\mathcal{F}_s$-measurable. Now we can use corollary 1.1.

$\square$

Analogously we can show that we have for the conditional expectation w.r.t. the σ-algebra $\mathcal{H}_s$

$$\int\limits_{\omega\in B} g\left(Z_T^z\right) Q\left(d\omega\right) =$$
$$= \int\limits_{\omega\in B} {}^{\omega^X}E^Y\left[E^X\left[g\left(X_{T-s}^{\alpha}\right)\right]\middle|\,\alpha = X_s^x(\omega^X)\middle|\,x = Y_{T-s}^{\beta}\right]\middle|\,\beta = Y_s^z(\omega^Y) Q\left(d\omega\right)$$

for every set $B \in \mathcal{H}_s$.

Since the set of measures ${}^{\omega^X}Q^Y$ is $\mathcal{F}_t^X$-measurable and $t < s$, then the following is true.

Proposition 1.5. *Let $t \leq s \leq T$. Then*

$$E\left[g\left(Z_T^z\right)/\mathcal{H}_s\right] =$$
$$= {}^{\omega^X}E^Y\left[E^X\left[g\left(X_{T-s}^{\alpha}\right)\right]\middle|\,\alpha = X_s^x(\omega^X)\middle|\,x = Y_{T-s}^{\beta}\right]\middle|\,\beta = Y_s^z(\omega^Y).$$

In the same way we can prove the analogous proposition for the conditional expectation of the "discounting" process.

Proposition 1.6.

$$\int\limits_{\omega\in B} \frac{Z_T^z}{\exp\left(\int\limits_s^T a\left(\overline{Z}_u^{\overline{z}}\right) ds\right)} Q\left(d\omega\right) =$$

$$= \int\limits_{\omega\in B} \left({}^{\omega^X}E^Y\left[E^X \frac{g\left(X_{T-s}^{\alpha}\right)\big|\,\alpha = X_s^x\left(\omega^X\right)\big|\,x = Y_{T-s}^{y}}{\exp\left(\int\limits_0^{T-s} a\left(\overline{X}_u^{\alpha}\right)\middle|\,\alpha = \overline{X}_s^x\left(\omega^X\right)\middle|\,x = \overline{Y}_u^{\overline{y}} du\right)}\right]\Bigg|\Bigg| \middle|\, y = Y_s^z\left(\omega^Y\right); \overline{y} = \overline{Y}_s^z\left(\omega^Y\right)\right) Q\left(d\omega\right)$$

for all $B \in \mathcal{F}_s$. Also

$$E\left[\frac{Z_T^z}{\exp\left(\int\limits_s^T a\left(\overline{Z}_u^{\overline{z}}\right) ds\right)}\middle|\, H_s\right]\left(\omega^X, \omega^Y\right) =$$

$$= {}^{\omega^X}E^Y\left[E^X\left[\frac{g\left(X^{\alpha}_{T-s}\right)\big|\,\alpha = X^x_s\left(\omega^X\right)\big|\,x = Y^y_{T-s}}{\exp\left(\int\limits_0^{T-s} a\left(\overline{X}^{\alpha}_u\right)\Big|\,\alpha = \overline{X}^x_s\left(\omega^X\right)\Big|\,x = \overline{Y}^{\overline{y}}_u du\right)}\right]\right]\Bigg|$$

$$\Big|y = Y^z_s\left(\omega^Y\right); \overline{y} = \overline{Y}^z_s\left(\omega^Y\right).$$

1.3 Infinitesimal Generator

Now we shall find the infinitesimal generator w.r.t. the measure Q for moments s, $t < s < T$.

Proposition 1.7.

$$A_s h(z) = E^X\left[{}^{X\omega}E^Y\left[\,A^X\left\{h\left(\alpha\right)\right\}\big|\,\alpha = X^x_s\big|\,x = Y^z_s\right]\right] +$$
$$+ E^X\left[{}^{X\omega}E^Y\left[{}^{X\omega}A^Y h\left(X^x_s\right)\Big|\,x = Y^z_s\right]\right].$$

Proof.

$$A^Q_s h(z) = \lim_{\varepsilon\to 0}\frac{E^Q\left[h\left(Z^z_{s+\varepsilon}\right)\right] - E^Q\left[h\left(Z^z_s\right)\right]}{\varepsilon} =$$

$$= \lim_{\varepsilon\to 0}\frac{E^X\left[{}^{\omega^X}E^Y\left[h\left(X^x_{s+\varepsilon}\right)\big|\,x = Y^z_{s+\varepsilon}\right]\right] - E^X\left[{}^{\omega^X}E^Y\left[h\left(X^x_s\right)\big|\,x = Y^z_s\right]\right]}{\varepsilon} =$$

$$= \lim_{\varepsilon\to 0}\left[\begin{array}{l}\dfrac{E^X\left[{}^{\omega^X}E^Y\left[h\left(X^x_{s+\varepsilon}\right)\big|\,x = Y^z_{s+\varepsilon}\right]\right] - E^X\left[{}^{\omega^X}E^Y\left[h\left(X^x_s\right)\big|\,x = Y^z_s\right]\right]}{\varepsilon}\pm\\ \pm\dfrac{E^X\left[{}^{\omega^X}E^Y\left[h\left(X^x_s\right)\big|\,x = Y^z_{s+\varepsilon}\right]\right]}{\varepsilon}\end{array}\right] =$$

$$= \lim_{\varepsilon\to 0}\left[\frac{E^X\left[{}^{\omega^X}E^Y\left[h\left(X^x_{s+\varepsilon}\right)\big|\,x = Y^z_{s+\varepsilon}\right]\right] - E^X\left[{}^{\omega^X}E^Y\left[h\left(X^x_s\right)\big|\,x = Y^z_{s+\varepsilon}\right]\right]}{\varepsilon}\right] +$$

$$+ \lim_{\varepsilon\to 0}\left[\frac{E^X\left[{}^{\omega^X}E^Y\left[h\left(X^x_s\right)\big|\,x = Y^z_{s+\varepsilon}\right]\right] - E^X\left[{}^{\omega^X}E^Y\left[h\left(X^x_s\right)\big|\,x = Y^z_s\right]\right]}{\varepsilon}\right] =$$

$$= L_1 + L_2.$$

We shall study separately the two parts again. Let us fix one of the measures ${}^{\omega^X}Q^Y$ as index and let all other measures have Radon-Nikodim derivatives $\xi\left(\omega^X, \omega^Y\right)$ w.r.t. it. Then

$$L_1 = \lim_{\varepsilon\to 0}\left[\frac{\begin{array}{l} E^X\left[{}^{\omega^X}E^Y\left[h\left(X^x_{s+\varepsilon}\right)\middle| x = Y^z_{s+\varepsilon}\right]\right] - \\ -E^X\left[{}^{\omega^X}E^Y\left[h\left(X^x_s\right)\middle| x = Y^z_{s+\varepsilon}\right]\right]\end{array}}{\varepsilon}\right] =$$

$$= \lim_{\varepsilon\to 0}\left[\frac{\begin{array}{l} E^X\left[\overline{E}^Y\left[\xi\left(\omega^X,\omega^Y\right)h\left(X^x_{s+\varepsilon}\right)\middle| x = Y^z_{s+\varepsilon}\right]\right] - \\ -E^X\left[\overline{E}^Y\left[\xi\left(\omega^X,\omega^Y\right)h\left(X^x_s\right)\middle| x = Y^z_{s+\varepsilon}\right]\right]\end{array}}{\varepsilon}\right] =$$

$$= \lim_{\varepsilon\to 0}\left[\frac{E^X\left[\overline{E}^Y\left[\xi\left(\omega^X,\omega^Y\right)\left(h\left(X^x_{s+\varepsilon}\right)-h\left(X^x_s\right)\right)\middle| x = Y^z_{s+\varepsilon}\right]\right]}{\varepsilon}\right] =$$

$$= \lim_{\varepsilon\to 0}\left[\frac{\overline{E}^Y\left[E^X\left[\xi\left(\omega^X,\omega^Y\right)\left(h\left(X^x_{s+\varepsilon}\right)-h\left(X^x_s\right)\right)\right]\middle| x = Y^z_{s+\varepsilon}\right]}{\varepsilon}\right]. \tag{8}$$

We have that for a fixed ω^Y, $\xi\left(\omega^X,\omega^Y\right)$ is $\mathcal{F}^X_t$-measurable and $h\left(X^x_{s+\varepsilon}\right) - h\left(X^x_s\right)$ is independent of $\mathcal{F}^X_s$ and therefore $h\left(X^x_{s+\varepsilon}\right) - h\left(X^x_s\right)$ is independent of $\xi\left(\omega^X,\omega^Y\right)$ (Note that $s > t$). Thus

$$\begin{array}{l} E^X\left[\left(h\left(X^x_{s+\varepsilon}\right)-h\left(X^x_s\right)\right)\xi\left(\omega^X,\omega^Y\right)\right] = \\ = E^X\left[h\left(X^x_{s+\varepsilon}\right)-h\left(X^x_s\right)\right]E^X\left[\xi\left(\omega^X,\omega^Y\right)\right]. \end{array}$$

Replacing in (8) :

$$L_1 = \lim_{\varepsilon\to 0}\left[\frac{\overline{E}^Y\left[E^X\left[\xi\left(\omega^X,\omega^Y\right)\left(h\left(X^x_{s+\varepsilon}\right)-h\left(X^x_s\right)\right)\right]\middle| x = Y^z_{s+\varepsilon}\right]}{\varepsilon}\right] =$$

$$= \lim_{\varepsilon\to 0}\left[\frac{\overline{E}^Y\left[E^X\left[\left(h\left(X^x_{s+\varepsilon}\right)-h\left(X^x_s\right)\right)\right]E^X\left[\xi\left(\omega^X,\omega^Y\right)\right]\middle| x = Y^z_{s+\varepsilon}\right]}{\varepsilon}\right] =$$

$$= \overline{E}^Y\left[E^X\left[A^X h\left(a\right)\middle| a = X^x_s\right]E^X\left[\xi\left(\omega^X,\omega^Y\right)\right]\middle| x = Y^z_s\right] =$$

$$= \overline{E}^Y\left[\int\limits_{\omega^X\in\Omega^X}\xi\left(\omega^X,\omega^Y\right){}^XQ\left(d\omega^X\right)E^X\left[A^X h\left(a\right)\middle| a = X^x_s\right]\middle| x = Y^z_s\right] =$$

$$= \overline{E}^Y\left[\int\limits_{\omega^X\in\Omega^X}\xi\left(\omega^X,\omega^Y\right)E^X\left[A^X h\left(a\right)\middle| a = X^x_s\right]\middle| x = Y^{z\,X}_sQ\left(d\omega^X\right)\right] =$$

$$
\begin{aligned}
&= \int\limits_{\omega^X \in \Omega^X} \overline{E}^Y \left[\xi\left(\omega^X, \omega^Y\right) E^X \left[A^X h(a) \middle| a = X_s^x \right] \middle| x = Y_s^z \right] Q\left(d\omega^X\right) = \\
&= \int\limits_{\omega^X \in \Omega^X} {}^{\omega^X} E^Y \left[E^X \left[A^X h(a) \middle| a = X_s^x \right] \middle| x = Y_s^z \right] Q\left(d\omega^X\right) = \\
&= E^X \left[{}^{\omega^X} E^Y \left[E^X \left[A^X h(a) \middle| a = X_s^x \right] \middle| x = Y_s^z \right] \right].
\end{aligned}
$$

We have for L_2

$$
\begin{aligned}
L_2 = \lim_{\varepsilon \to 0} & \frac{E^X \left[{}^{\omega^X} E^Y \left[h\left(X_s^x\right) \middle| x = Y_{s+\varepsilon}^z \right] \right] - E^X \left[{}^{\omega^X} E^Y \left[h\left(X_s^x\right) \middle| x = Y_s^z \right] \right]}{\varepsilon} = \\
&= \lim_{\varepsilon \to 0} E^X \left[\frac{{}^{\omega^X} E^Y \left[h\left(X_s^x\right) \middle| x = Y_{s+\varepsilon}^z \right] - {}^{\omega^X} E^Y \left[h\left(X_s^x\right) \middle| x = Y_s^z \right]}{\varepsilon} \right] = \\
&= \lim_{\varepsilon \to 0} E^X \left[{}^{\omega^X} E^Y \left[\frac{{}^{\omega^X} E^Y \left[h\left(X_s^x\right) \middle| x = Y_\varepsilon^z \right] - {}^{\omega^X} E^Y \left[h\left(X_s^x\right) \right]}{\varepsilon} \middle| x = Y_s^z \right] \right] = \\
&= E^X \left[{}^{\omega^X} E^Y \left[{}^{\omega^X} A^Y \left\{ h\left(X_s^x\right) \right\} \middle| x = Y_s^z \right] \right]
\end{aligned}
$$

or finally

$$
\begin{aligned}
A_s h(z) &= L_1 + L_2 = \\
&= E^X \left[{}^{X\omega} E^Y \left[A^X \left\{ h(\alpha) \right\} \middle| \alpha = X_s^x \middle| x = Y_s^z \right] \right] + \\
&+ E^X \left[{}^{X\omega} E^Y \left[{}^{X\omega} A^Y h\left(X_s^x\right) \middle| x = Y_s^z \right] \right].
\end{aligned}
$$

□

1.4 Radon-Nikodim derivative

Now we shall find the Radon-Nikodim derivative between two measures, which are constructed in the above way.

Let the measures be Q and $\overline{Q}$ and the corresponding Radon-Nikodim derivatives of the measures Q^X and ${}^{\omega^X} Q^Y$ w.r.t. $\overline{Q}^X$ and ${}^{\omega^X}\overline{Q}^Y$ be $\xi^X\left(\omega^X\right)$ and ${}^{\omega^X}\xi^Y\left(\omega^Y\right)$. Let $\xi\left(\omega^X, \omega^Y\right)$ be the Radon-Nikodim derivative of the measure Q w.r.t. $\overline{Q}$.

Proposition 1.8. *The relation between the Radon-Nikodim derivatives is*

$$\xi\left(\omega^X, \omega^Y\right) = \xi\left(\omega^X\right) {}^{\omega^X}\xi^Y\left(\omega^Y\right).$$

Proof. Let us examine the form of the expectation in the proposition 1.3

$$\begin{aligned} E^{Q}\left(h(Z_s^z)\right) &= E^{Q,X}\left[{}^{\omega^X}E^{Q,Y}\left[h(X_s^x)\right| x = Y_s^z]\right] = \\ &= E^{\overline{Q},X}\left[\xi\left(\omega^X\right){}^{\omega^X}E^{Q,Y}\left[h(X_s^x)\right| x = Y_s^z]\right] = \\ &= E^{\overline{Q},X}\left[\xi\left(\omega^X\right)E^{\overline{Q},Y}\left[{}^{\omega^X}\xi^Y\left(\omega^Y\right)h(X_s^x)\right| x = Y_s^z\right]\right] = \\ &= E^{\overline{Q},X}\left[E^{\overline{Q},Y}\left[\xi\left(\omega^X\right){}^{\omega^X}\xi^Y\left(\omega^Y\right)h(X_s^x)\right| x = Y_s^z\right]\right]. \end{aligned}$$

Since

$$\begin{aligned} E\left[\xi\left(\omega^X\right){}^{\omega^X}\xi^Y\left(\omega^Y\right)\right] &= E^X\left[E^Y\left[\xi\left(\omega^X\right){}^{\omega^X}\xi^Y\left(\omega^Y\right)\right]\right] = \\ &= E^X\left[\xi\left(\omega^X\right)E^Y\left[{}^{\omega^X}\xi^Y\left(\omega^Y\right)\right]\right] = \\ &= E^X\left[\xi\left(\omega^X\right)\right] = 1, \end{aligned}$$

then we see that

$$\xi\left(\omega^X,\omega^Y\right) = \xi\left(\omega^X\right){}^{\omega^X}\xi^Y\left(\omega^Y\right).$$

□

2 Pseudomartingale measure. Construction

Definition 2.1. *We call a pseudomartingale measure for the process Z_t a class of measures Q^t, equivalent to the natural measure P, for which the condition*

$$E^{Q^t}\left[Z_T\right|\mathcal{F}_t] = Z_t \quad \forall t \le T$$

is satisfied.

As we can see from the definition, we do not consider a constant measure, but different measures for every moment. In this section we shall give a method to construct these measures.

1. Let Q^X be any equivalent measure for the process X_t^x.
2. Let $h(t,x)$ be a function $h(t,x) = E^{Q^X}\left[g\left(X_t^\alpha\right)\right]\Big| \alpha = X_s^x\left(\omega^x\right)$ for a fixed moment s and a state ω^X. Let also ${}^{\omega^X}Q_s^Y$ be an equivalent measure for the process Y, for which $h\left(t, Y_t^z\right)$ is a martingale. Note that ${}^{\omega^X}Q_s^Y$ is $\mathcal{F}_s^X$-measurable. Denote that Q^X may be dependent on s without any restrictions.
3. Using the reasons in section 1 we define measures Q_s, equivalent to the natural measure P.

We shall show that this class of measures is a pseudomartingale measure.

Proposition 2.1. *The class of measures Q_s is a pseudomartingale measure for the process $g(Z)$.*

Proof. Using the propositions for the conditional expectation (proposition 1.4 , corollaries 1.1 and 1.2) we have for arbitrary an $A \in \mathcal{F}_s$

$$\int_A E^{Q_s}\left[g(Z^z_{t+s})/\mathcal{F}_s\right]\left(\omega^X,\omega^Y\right)Q\,(d\omega) =$$
$$= \int_A E^{{}^{\omega^X}Q^Y_s}\left[E^{Q^X}\left[g\left(X^\alpha_t\right)\right]\Big|\,\alpha = X^x_s(\omega^X)\Big|\,x = Y^\beta_t\right]\Big|\,\beta = Y^z_s(\omega^Y)Q\,(d\omega)\,.$$

We use the definition of ${}^{\omega^X}Q^Y_s$, i.e. it is a martingale measure for the process $h\,(t, Y^z_t)$, where $h\,(t,x) = E^{Q^X}\left[g\left(X^\alpha_t\right)\right]\Big|\,\alpha = X^x_s\left(\omega^x\right)$, and therefore

$$\begin{aligned} g\left(X^\beta_s\left(\omega^x\right)\right) = h\,(0,\beta) &= \\ &= E^{{}^{\omega^x}Q^Y_s}\left[h\left(t, Y^\beta_t\right)\right] = \\ &= E^{{}^{\omega^x}Q^Y_s}\left[E^{Q^X}\left[g\left(X^\alpha_t\right)\right]\Big|\,\alpha = X^x_s\left(\omega^x\right)\big|\,x = Y^\beta_t\right]. \end{aligned}$$

Thus

$$\begin{aligned} &E^{{}^{\omega^X}Q^Y_s}\left[E^{Q^X}\left[g\left(X^\alpha_t\right)\right]\Big|\,\alpha = X^x_s(\omega^X)\Big|\,x = Y^\beta_t\right]\Big|\,\beta = Y^z_s(\omega^Y) = \\ &= g\left(X^\beta_s(\omega^X)\right)\big|\,\beta = Y^z_s(\omega^Y) \end{aligned}$$

is $\mathcal{F}_s$-measurable and from the corollary 1.1 we get

$$\begin{aligned} &E^{Q_s}\left[g(Z^z_{t+s})/\mathcal{F}_s\right]\left(\omega^X,\omega^Y\right) = \\ &= E^{{}^{\omega^X}Q^Y_s}\left[E^{Q^X}\left[g\left(X^\alpha_t\right)\right]\Big|\,\alpha = X^x_s(\omega^X)\Big|\,x = Y^\beta_t\right]\Big|\,\beta = Y^z_s(\omega^Y) = \\ &= g\left(X^x_s(\omega^X)\right)\big|\,x = Y^z_s(\omega^Y), \end{aligned}$$

which shows that Q_s is a pseudomartingale measure for the process Z.

□

3 Agreement condition. Martingale measure

Now we shall discuss a naturally appearing agreement condition between the measures of class Q_s and we shall give an example for constructing a class of measures, which satisfy this condition.

Condition 3.1 (Agreement condition). *Let $s<t$ and ${}_{s,t}\xi$ be a Radon-Nikodim derivative between the measures Q_s and Q_t. We shall say that the agreement condition is satisfied if ${}_{s,t}\xi$ is $\mathcal{F}_t$-measurable.*

Let also ${}_{s,t}\xi_l = E\left[{}_{s,t}\xi/\mathcal{F}_l\right]$.

For three moments $s<t<T$ we have (following from the pseudomartingality):

$$E^{Q_t}\left(Z_T/\mathcal{F}_t\right) = Z_t \tag{9}$$

$$E^{Q_s}\left(Z_T/\mathcal{F}_s\right) = Z_s \tag{10}$$

$$E^{Q_s}\left(Z_t/\mathcal{F}_s\right) = Z_s. \tag{11}$$

Therefore:

$$\begin{aligned} E^{Q_s}\left(Z_T/\mathcal{F}_s\right) &= E^{Q_s}\left[Z_t/\mathcal{F}_s\right] = \\ &= E^{Q_s}\left[E^{Q_t}\left(Z_T/\mathcal{F}_t\right)/\mathcal{F}_s\right] = \\ &= E^{Q_s}\left[\frac{E^{Q_s}\left({}_{s,t}\xi Z_T/\mathcal{F}_t\right)}{E^{Q_s}\left({}_{s,t}\xi/\mathcal{F}_t\right)}/\mathcal{F}_s\right]. \end{aligned}$$

If the agreement condition is satisfied, i.e. ${}_{s,t}\xi$ is $\mathcal{F}_t$-measurable, then:

$$\begin{aligned} E^{Q_s}\left[\frac{E^{Q_s}\left({}_{s,t}\xi Z_T/\mathcal{F}_t\right)}{E^{Q_s}\left({}_{s,t}\xi/\mathcal{F}_t\right)}/\mathcal{F}_s\right] &= E^{Q_s}\left(\frac{{}_{s,t}\xi E^{Q_s}\left(Z_T/\mathcal{F}_t\right)}{{}_{s,t}\xi}/\mathcal{F}_s\right) = \\ &= E^{Q_s}\left(Z_T/\mathcal{F}_s\right). \end{aligned}$$

So the agreement condition guarantees that there is no contradiction between (9), (10) and (11).

Proposition 3.1. *If the agreement condition is satisfied, then*

$$E^{Q_s}\left[E^{Q_t}\left(\eta/\mathcal{F}_t\right)/\mathcal{F}_s\right] = E^{Q_s}\left(\eta/\mathcal{F}_s\right)$$

for an arbitrary $\mathcal{F}_T$ -measurable η.

Proof. Using the fact that ${}_{s,t}\xi$ is $\mathcal{F}_t$-measurable (i.e. the agreement condition is satisfied) we have

$$\begin{aligned} E^{Q_s}\left[E^{Q_t}\left(\eta/\mathcal{F}_t\right)/\mathcal{F}_s\right] &= E^{Q_s}\left[\frac{E^{Q_s}\left({}_{s,t}\xi\eta/\mathcal{F}_t\right)}{E^{Q_s}\left({}_{s,t}\xi/\mathcal{F}_t\right)}/\mathcal{F}_s\right] = \\ &= E^{Q_s}\left[\frac{E^{Q_s}\left(\eta/\mathcal{F}_t\right){}_{s,t}\xi}{{}_{s,t}\xi}/\mathcal{F}_s\right] = \\ &= E^{Q_s}\left[E^{Q_s}\left(\eta/\mathcal{F}_t\right)/\mathcal{F}_s\right] \\ &= E^{Q_s}\left[\eta/\mathcal{F}_s\right]. \end{aligned}$$

□

Proposition 3.2. *If for three arbitrary moments $l < s < t$, ${}_{s,t}\xi$ is $\mathcal{F}_t$-measurable and ${}_{l,s}\xi$ is $\mathcal{F}_s$-measurable, then ${}_{l,t}\xi$ is $\mathcal{F}_t$-measurable. (${}_{l,t}\xi$, ${}_{s,t}\xi$ and ${}_{l,s}\xi$ are the corresponding Radon-Nikodim derivatives)*

Proof. Let η be an arbitrary $\mathcal{F}$-measurable random variable. From the fact that ${}_{l,s}\xi$ is $\mathcal{F}_s$-measurable we have that

$$\begin{aligned} E^{Q_l}\left[\eta/\mathcal{F}_l\right] &= E^{Q_l}\left[E^{Q_l}\left(\eta/\mathcal{F}_s\right)/\mathcal{F}_l\right] = \\ &= E^{Q_l}\left[\frac{E^{Q_l}\left(\eta/\mathcal{F}_s\right){}_{l,s}\xi}{{}_{l,s}\xi}/\mathcal{F}_l\right] = \\ &= E^{Q_l}\left[\frac{E^{Q_l}\left({}_{l,s}\xi\eta/\mathcal{F}_s\right)}{E^{Q_l}\left({}_{l,s}\xi/\mathcal{F}_s\right)}/\mathcal{F}_l\right] = \\ &= E^{Q_l}\left[E^{Q_s}\left(\eta/\mathcal{F}_s\right)/\mathcal{F}_l\right]. \end{aligned} \tag{12}$$

On the other hand from the fact that ${}_{s,t}\xi$ is $\mathcal{F}_t$-measurable, we have

$$\begin{aligned} E^{Q_s}\left[\eta/\mathcal{F}_s\right] &= E^{Q_s}\left[E^{Q_s}\left(\eta/\mathcal{F}_t\right)/\mathcal{F}_s\right] = \\ &= E^{Q_s}\left[\frac{E^{Q_s}\left(\eta/\mathcal{F}_t\right){}_{s,t}\xi}{{}_{s,t}\xi}/\mathcal{F}_s\right] = \\ &= E^{Q_s}\left[\frac{E^{Q_s}\left({}_{s,t}\xi\eta/\mathcal{F}_t\right)}{E^{Q_s}\left({}_{s,t}\xi/\mathcal{F}_t\right)}/\mathcal{F}_s\right] = \\ &= E^{Q_s}\left[E^{Q_t}\left(\eta/\mathcal{F}_t\right)/\mathcal{F}_s\right]. \end{aligned} \tag{13}$$

Replacing (13) in (12) we reach

$$E^{Q_l}\left[\eta/\mathcal{F}_l\right] = E^{Q_l}\left[E^{Q_s}\left[E^{Q_t}\left(\eta/\mathcal{F}_t\right)/\mathcal{F}_s\right]/\mathcal{F}_l\right].$$

Using that $E^{Q_l}\left[\mu/\mathcal{F}_l\right] = E^{Q_l}\left[E^{Q_s}\left(\mu/\mathcal{F}_s\right)/\mathcal{F}_l\right]$ for $\mu = E^{Q_t}\left(\eta/\mathcal{F}_t\right)$ we reach

$$\begin{aligned} E^{Q_l}\left[\eta/\mathcal{F}_l\right] &= E^{Q_l}\left[E^{Q_s}\left[E^{Q_t}\left(\eta/\mathcal{F}_t\right)/\mathcal{F}_s\right]/\mathcal{F}_l\right] = \\ &= E^{Q_l}\left[E^{Q_s}\left[\mu/\mathcal{F}_s\right]/\mathcal{F}_l\right] = \\ &= E^{Q_l}\left[\mu/\mathcal{F}_l\right] = \\ &= E^{Q_l}\left[E^{Q_t}\left(\eta/\mathcal{F}_t\right)/\mathcal{F}_l\right] = \\ &= E^{Q_l}\left[\frac{E^{Q_l}\left({}_{l,t}\xi\eta/\mathcal{F}_t\right)}{E^{Q_l}\left({}_{l,t}\xi/\mathcal{F}_t\right)}/\mathcal{F}_l\right]. \end{aligned}$$

Since this is true for an arbitrary $\mathcal{F}$-measurable η, then ${}_{l,t}\xi$ is $\mathcal{F}_t$-measurable, which we wanted to show.

□

Now we shall show in several steps that, with some modifications, the pseudomartingale measure, defined in the previous section, satisfies the agreement condition.

1. Separate the time interval $[0,T]$ into n intervals $0 = t_1 < t_2 < ... < t_{n-1} < t_n = T$.

2. Construct a measure $Q_{t_{n-1}}$, such that $E^{Q_{t_{n-1}}}\left(Z_T/\mathcal{F}_{t_{n-1}}\right) = Z_{t_{n-1}}$.

3. Construct a measure $Q_{t_{n-2}}$, such that $E^{Q_{t_{n-2}}}\left(Z_{t_{n-1}}/\mathcal{F}_{t_{n-2}}\right) = Z_{t_{n-2}}$ and $E^{Q_{t_{n-2}}}\left(Z_T/\mathcal{F}_{t_{n-2}}\right) = Z_{t_{n-2}}$.

4. Let $Q^X_{t_{n-1}}$, ${}^{\omega^X}Q^Y_{t_{n-1}}$, $Q^X_{t_{n-2}}$ and ${}^{\omega^X}Q^Y_{t_{n-2}}$ be the corresponding measures with the help of which the measures $Q_{t_{n-1}}$ and $Q_{t_{n-2}}$ are constructed. Let

$$\begin{gathered}\xi^X_{n-2,n-1}\left(\omega^X\right)\\ {}^{\omega^X}\xi^Y_{n-2,n-1}\left(\omega^Y\right)\end{gathered}$$

be the corresponding Radon-Nikodim derivatives of the measures $Q^X_{t_{n-2}}$ and ${}^{\omega^X}Q^Y_{t_{n-2}}$ w.r.t. $Q^X_{t_{n-1}}$ and ${}^{\omega^X}Q^Y_{t_{n-1}}$. Note that for every fixed ω^Y ${}^{\omega^X}\xi^Y\left(\omega^Y\right)$ is $\mathcal{F}^X_{t_{n-1}}$-measurable, because ${}^{\omega^X}Q^Y_{t_{n-2}}$ is $\mathcal{F}^X_{t_{n-2}}$-measurable and ${}^{\omega^X}Q^Y_{t_{n-1}}$ is $F^X_{t_{n-1}}$-measurable. Let $\xi_{n-2,n-1}\left(\omega^X,\omega^Y\right)$ be the Radon-Nikodim derivative between the measures $Q_{t_{n-1}}$ and $Q_{t_{n-2}}$. We know from the proposition 1.8 that

$$\xi_{n-2,n-1}\left(\omega^X,\omega^Y\right) = \xi^X_{n-2,n-1}\left(\omega^X\right){}^{\omega^X}\xi^X_{n-2,n-1}\left(\omega^Y\right).$$

Let us examine the measures $\overline{Q}^X_{t_{n-2}}$ and ${}^{\omega^X}\overline{Q}^Y_{t_{n-2}}$ for which the Radon-Nikodim derivatives w.r.t. the measures $Q^X_{t_{n-1}}$ and ${}^{\omega^X}Q^Y_{t_{n-1}}$ are respectively

$$\begin{aligned}&\overline{\xi}^X_{n-2,n-1}\left(\omega^X\right) = E^{Q^X_{t_{n-1}}}\left[\xi^X_{n-2,n-1}\left(\omega^X\right)/\mathcal{F}^X_{t_{n-1}}\right]\\ &{}^{\omega^X}\overline{\xi}^Y_{n-2,n-1}\left(\omega^Y\right) = E^{{}^{\omega^X}Q^Y_{t_{n-1}}}\left[{}^{\omega^X}\xi^Y_{n-2,n-1}\left(\omega^Y\right)\middle|\mathcal{F}^Y_{t_{n-1}}\right].\end{aligned}$$

Note that the set ${}^{\omega^X}\overline{Q}^Y_{t_{n-2}}$ is also $\mathcal{F}^X_{t_{n-2}}$-measurable. Let $\overline{Q}_{t_{n-2}}$ be the measure constructed by the method in the section 1. We have for the Radon-Nikodim derivative between the measures $\overline{Q}_{t_{n-2}}$ and $Q_{t_{n-1}}$

$$\begin{aligned}&\overline{\xi}_{n-2,n-1}\left(\omega^X,\omega^Y\right) = \overline{\xi}^X_{n-2,n-1}\left(\omega^X\right){}^{\omega^X}\overline{\xi}^Y_{n-2,n-1}\left(\omega^Y\right) =\\ &= E^{Q^X_{t_{n-1}}}\left[\xi^X_{n-2,n-1}\left(\omega^X\right)/\mathcal{F}^X_{t_{n-1}}\right]E^{{}^{\omega^X}Q^Y_{t_{n-1}}}\left[{}^{\omega^X}\xi^Y_{n-2,n-1}\left(\omega^Y\right)\middle|\mathcal{F}^Y_{t_{n-1}}\right].\end{aligned}$$

We use that the set ${}^{\omega^X}\overline{Q}^Y_{t_{n-2}}$ is $\mathcal{F}^X_{t_{n-2}}$-measurable and ${}^{\omega^X}\xi^Y_{n-2,n-1}\left(\omega^Y\right)$ is $\mathcal{F}^X_{t_{n-1}}$-measurable for every fixed ω^Y, to obtain

$$
\begin{aligned}
&\overline{\xi}_{n-2,n-1}\left(\omega^X,\omega^Y\right) = \\
&= E^{Q^X_{t_{n-1}}}\left[\xi^X_{n-2,n-1}\left(\omega^X\right)/\mathcal{F}^X_{t_{n-1}}\right] E^{{}^{\omega^X}Q^Y_{t_{n-1}}}\left[{}^{\omega^X}\xi^Y_{n-2,n-1}\left(\omega^Y\right)/\mathcal{F}^Y_{t_{n-1}}\right] = \\
&= E^{Q^X_{t_{n-1}}}\left[\xi^X_{n-2,n-1}\left(\omega^X\right) E^{{}^{\omega^X}Q^Y_{t_{n-1}}}\left[{}^{\omega^X}\xi^Y_{n-2,n-1}\left(\omega^Y\right)/\mathcal{F}^Y_{t_{n-1}}\right]/\mathcal{F}^X_{t_{n-1}}\right] = \\
&= E^{Q^X_{t_{n-1}}}\left[E^{{}^{\omega^X}Q^Y_{t_{n-1}}}\left[\xi^X_{n-2,n-1}\left(\omega^X\right){}^{\omega^X}\xi^Y_{n-2,n-1}\left(\omega^Y\right)/\mathcal{F}^Y_{t_{n-1}}\right]/\mathcal{F}^X_{t_{n-1}}\right] = \\
&= E^{Q_{t_{n-1}}}\left[\xi^X_{n-2,n-1}\left(\omega^X\right){}^{\omega^X}\xi^Y_{n-2,n-1}\left(\omega^Y\right)/\mathcal{F}_{t_{n-1}}\right] = \\
&= E^{Q_{t_{n-1}}}\left[\xi_{n-2,n-1}\left(\omega^X,\omega^Y\right)/\mathcal{F}_{t_{n-1}}\right].
\end{aligned}
$$

Note that $\overline{\xi}_{n-2,n-1}\left(\omega^X,\omega^Y\right)$ is $\mathcal{F}_{t_{n-1}}$-adapted. Now we can show that

$$
\begin{aligned}
&E^{\overline{Q}_{t_{n-2}}}\left(Z_{t_{n-1}}/\mathcal{F}_{t_{n-2}}\right) = Z_{t_{n-2}} \\
&E^{\overline{Q}_{t_{n-2}}}\left(Z_T/\mathcal{F}_{t_{n-2}}\right) = Z_{t_{n-2}}.
\end{aligned}
$$

We have

$$
\begin{aligned}
&E^{\overline{Q}_{t_{n-2}}}\left(Z_{t_{n-1}}/\mathcal{F}_{t_{n-2}}\right) = \frac{E^{Q_{t_{n-1}}}\left(\overline{\xi}_{n-2,n-1}Z_{t_{n-1}}/\mathcal{F}_{t_{n-2}}\right)}{E^{Q_{t_{n-1}}}\left(\overline{\xi}_{n-2,n-1}/\mathcal{F}_{t_{n-2}}\right)} = \\
&= \frac{E^{Q_{t_{n-1}}}\left(E^{Q_{t_{n-1}}}\left[\xi_{n-2,n-1}\left(\omega^X,\omega^Y\right)/\mathcal{F}_{t_{n-1}}\right]Z_{t_{n-1}}/\mathcal{F}_{t_{n-2}}\right)}{E^{Q_{t_{n-1}}}\left(E^{Q_{t_{n-1}}}\left[\xi_{n-2,n-1}\left(\omega^X,\omega^Y\right)/\mathcal{F}_{t_{n-1}}\right]/\mathcal{F}_{t_{n-2}}\right)} = \\
&= \frac{E^{Q_{t_{n-1}}}\left(E^{Q_{t_{n-1}}}\left[\xi_{n-2,n-1}\left(\omega^X,\omega^Y\right)Z_{t_{n-1}}/\mathcal{F}_{t_{n-1}}\right]/\mathcal{F}_{t_{n-2}}\right)}{E^{Q_{t_{n-1}}}\left(E^{Q_{t_{n-1}}}\left[\xi_{n-2,n-1}\left(\omega^X,\omega^Y\right)/\mathcal{F}_{t_{n-1}}\right]/\mathcal{F}_{t_{n-2}}\right)} = \\
&= \frac{E^{Q_{t_{n-1}}}\left(\xi_{n-2,n-1}\left(\omega^X,\omega^Y\right)Z_{t_{n-1}}/\mathcal{F}_{t_{n-2}}\right)}{E^{Q_{t_{n-1}}}\left(\xi_{n-2,n-1}\left(\omega^X,\omega^Y\right)/\mathcal{F}_{t_{n-2}}\right)} = \\
&= E^{Q_{t_{n-2}}}\left(Z_{t_{n-1}}/\mathcal{F}_{t_{n-2}}\right) = Z_{t_{n-2}}
\end{aligned}
$$

and

$$
\begin{aligned}
&E^{\overline{Q}_{t_{n-2}}}\left(Z_T/\mathcal{F}_{t_{n-2}}\right) = E^{\overline{Q}_{t_{n-2}}}\left(E^{\overline{Q}_{t_{n-2}}}\left(Z_T|\,\mathcal{F}_{t_{n-1}}\right)/\mathcal{F}_{t_{n-2}}\right) = \\
&= E^{\overline{Q}_{t_{n-2}}}\left(\frac{E^{Q_{t_{n-1}}}\left(\overline{\xi}_{n-2,n-1}Z_T|\,\mathcal{F}_{t_{n-1}}\right)}{E^{Q_{t_{n-1}}}\left(\overline{\xi}_{n-2,n-1}|\,\mathcal{F}_{t_{n-1}}\right)}/\mathcal{F}_{t_{n-2}}\right) =
\end{aligned}
$$

$$= E^{\overline{Q}_{t_{n-2}}} \left(\frac{\overline{\xi}_{n-2,n-1} E^{Q_{t_{n-1}}} \left(Z_T \middle| \mathcal{F}_{t_{n-1}} \right)}{\overline{\xi}_{n-2,n-1}} / \mathcal{F}_{t_{n-2}} \right) =$$
$$= E^{\overline{Q}_{t_{n-2}}} \left(Z_{t_{n-1}} / \mathcal{F}_{t_{n-2}} \right) = Z_{t_{n-2}}.$$

In this way we modify the measure $Q_{t_{n-2}}$ to $\overline{Q}_{t_{n-2}}$.

5. Analogously we continue back in the time, i.e. in each step we construct a measure Q_{t_i}, for which we want $E^{Q_{t_i}} \left(Z_{t_j} / \mathcal{F}_{t_i} \right) = Z_{t_i}$ for $i < j$. We construct the measures $\overline{Q}_{t_i}^X$ and ${}^{\omega^X}\overline{Q}_{t_i}^Y$ with the Radon-Nikodim derivatives w.r.t. $Q_{t_{i+1}}^X$ and ${}^{\omega^X}Q_{t_{i+1}}^Y$ respectively

$$\overline{\xi}_{i,i+1}^X \left(\omega^X \right) = E^{Q_{t_{i+1}}^X} \left[\xi_{i,i+1}^X \left(\omega^X \right) / \mathcal{F}_{t_{i+1}}^X \right]$$
$${}^{\omega^X}\overline{\xi}_{i,i+1}^Y \left(\omega^Y \right) = E^{{}^{\omega^X}Q_{t_{i+1}}^Y} \left[{}^{\omega^X}\xi_{i,i+1}^Y \left(\omega^Y \right) \middle| \mathcal{F}_{t_{i+1}}^Y \right].$$

We modify the measure Q_{t_i} into $\overline{Q}_{t_i}$, which is constructed by the measures $\overline{Q}_{t_i}^X$ and ${}^{\omega^X}\overline{Q}_{t_i}^Y$. We can prove in the same way that the Radon-Nikodim derivative between the measures $\overline{Q}_{t_i}$ and $Q_{t_{i+1}}$ is

$$\overline{\xi}_{i,i+1} \left(\omega^X, \omega^Y \right) = E^{Q_{t_{i+1}}} \left[\xi_{i,i+1} \left(\omega^X, \omega^Y \right) / \mathcal{F}_{t_{i+1}} \right],$$

i.e. it is $\mathcal{F}_{t_{i+1}}$-adapted. We can use the proposition 3.2 to see that $\overline{\xi}_{i,j} \left(\omega^X, \omega^Y \right)$ (the Radon-Nikodim derivative between the measures $\overline{Q}_{t_i}$ and Q_{t_j}) is $\mathcal{F}_{t_j}$-adapted. Analogously we prove that

$$E^{\overline{Q}_{t_i}} \left(Z_{t_j} / \mathcal{F}_{t_i} \right) = Z_{t_i} \forall j > i.$$

So we see that the modified measure $\overline{Q}_{t_i}$ satisfies the wanted conditions.

6. We must leave the length of the time intervals to goes to zero to finish the construction of the wanted measure.

We must note that the so constructed set of measures Q_s satisfies the agreement condition and something more – it is constructed by the method in the section 1 and the corresponding set ${}^{\omega^X}Q_s^Y$ is $\mathcal{F}_s^X$-measurable.

Now we shall give an important result, which is the existence of a martingale measure and a method to construct it.

Proposition 3.3. *If Q_s is a pseudomartingale measure and the agreement condition is satisfied, then for every $l < s < t$ we have $E^{Q_l} \left[Z_t / \mathcal{F}_s \right] = Z_s$.*

Proof. Since the agreement condition is satisfied and therefore ${}_{l,s}\xi$ is $\mathcal{F}_s$-measurable, then

$$\begin{aligned} E^{Q_l}\left[Z_t/\mathcal{F}_s\right] &= \frac{E^{Q_s}\left[{}_{l,s}\xi Z_t/\mathcal{F}_s\right]}{E^{Q_s}\left[{}_{l,s}\xi/\mathcal{F}_s\right]} = \\ &= \frac{{}_{l,s}\xi E^{Q_s}\left[Z_t/\mathcal{F}_s\right]}{{}_{l,s}\xi} = \\ &= E^{Q_s}\left(Z_t/\mathcal{F}_s\right) = \\ &= Z_s. \end{aligned}$$

□

Corollary 3.1. *Q_0 is a martingale measure for the process Z.*

Proof. Let $s < t$. Then $E^{Q_0}\left[Z_t/\mathcal{F}_s\right] = E^{Q_s}\left[Z_t/\mathcal{F}_s\right] = Z_s$.

□

4 Martingale measure

We shall show in this section when a measure is a martingale measure. As we showed in the previous section, the corollary 3.1 is true. In this way we construct a martingale measure for the Ito-Levi processes by using the pseudomartingale measure Q_t. From now on we shall study only measures, which are constructed with the help of the generalized probability measures in the section 1.

Suppose now that the measure Q is a martingale for the process $g(Z_t^z)$. Let the corresponding measures be Q_0^X and ${}^{\omega^X}Q_0^Y$. Since we shall work only with these measures, we shall drop the special symbols in the expectation. Let us suppose that

$${}^{\omega^X}E^Y\left[E^X\left[g\left(X_t^\alpha\right)\right]\big|\,\alpha = X_s^x(\omega^X)\big|\,x = Y_t^\beta\right]\Big|\,\beta = Y_s^z(\omega^Y)$$

is $\mathcal{F}_s$-measurable (or let us examine the conditional expectation w.r.t. the σ-algebra $\mathcal{H}$, generated from X_s^x for all x, and Y_s) and therefore

$$\begin{aligned} g(Z_s^z) &= g\left(X_s^x\right)\big|\,x = Y_s^z = \\ &= E^s\left[g(Z_{t+s}^z)/\mathcal{F}_s\right]\left(\omega^X,\omega^Y\right) = \\ &= {}^{\omega^X}E^Y\left[E^X\left[g\left(X_t^\alpha\right)\right]\big|\,\alpha = X_s^x(\omega^X)\big|\,x = Y_t^\beta\right]\Big|\,\beta = Y_s^z(\omega^Y). \end{aligned}$$

Thus

$$g\left(X_s^z\right) = {}^{\omega^X}E^Y\left[E^X\left[g\left(X_t^\alpha\right)\right]\big|\,\alpha = X_s^x(\omega^X)\big|\,x = Y_t^z\right].$$

Let us fix ω^X and a moment s and examine $h\left(t,x\right) = E^X\left[g\left(X_t^\alpha\right)\right]\big|\,\alpha = X_s^x(\omega^X)$. We see that ${}^{\omega^X}E^Y\left[h\left(t,Y_t^z\right)\right] = g\left(X_s^z\right)$, i.e. ${}^{\omega^X}E^Y\left[h\left(t,Y_t^z\right)\right]$ is constant for an

arbitrary t. We can use [10, proposition 8.3 points 2 and 3] to see that $h(t, Y_t^z)$ is a martingale. So

$$h(t, Y_t^z) = E^X [g(X_t^\alpha)] \big| \alpha = X_s^x(\omega^X) \big| x = Y_t^z$$

is a martingale w.r.t. the measure ${}^{\omega^X}Q^Y$. So we see that if the measure is a martingale and is derived by a generalized measure (section 1), then we must follow the procedure in section 2. This proves the fact that this procedure is the unique one for constructing of martingale measures of this type.

References

[1] J. Bertoin, *Levy Processes*, Cambridge University Press, 1976

[2] S. I. Boyarchenko and S. Z. Levendorskii, *Non-Gaussian Merton-Black-Scholes Theory*,World Scientific, New Jersey, London, Singapore, Hong Kong, 2002

[3] R. Cont, P. Tankov, *Financial modeling with jump processes*, Chapman & Hall 2004

[4] J. Lamperti, *Stochastic processes*, Springer-Verlag 1977

[5] L. Nirenberg, *Lectures on linear partial differential equations*, American Mathematical Society, 1973

[6] B. Oksendal,*Stochastic differential equations: an introduction with aplication*, Springer-Verlag 1998

[7] S. Rachev, S. Mittnik,*Stable Paretian Models in Finance*, John Wiley&Sons Ltd, 2000

[8] G. Samorodnitsky and M. S. Taqqu, *Stable Non-Gaussian Random Processes*, Chapman & Hall, New York, London

[9] K. Sato, *Levy processes and infinitely divisible distributions*, Cambridge University Press 1999

[10] T.S. Zaevski, *Ito-Levy processes*

JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS,VOL.5,NO.4,323-329,2007,COPYRIGHT 2007 EUDOXUS PRESS,LLC

An Adomian Decomposition Method for Predator-Prey Model Equations

Elias Deeba, Jeong-Mi Yoon, Shishen Xie
Department of Computer & Mathematical Sciences,
University of Houston-Downtown,
One Main Street, Houston, TX 77002, USA
Yoonj@uhd.edu, Xies@uhd.edu

Abstract: In this paper we analyze the Predator-Prey Model using an Adomian Decomposition Method (ADM) and then compare the results with the 4^{th}-order Runge-Kutta Method numerically.

Key Words: Predator-Prey Model, Adomian Decomposition Method.

Classification: 34E05, 37M05, 41A10

1 Introduction

Recently, S. Olek and J. T. Edwards, et. al. used the Adomian Decomposition Method (ADM) for solving the Predator-Prey Model and comparing it with numerical RK 4th-order Method (See [7] & [4]). Here we choose the coupled first order Predator-Prey Model which was used by J. T. Edwards, et. al. ('99) [4] and show the improvement of the result of ADM by controlling the computational method for nonlinear part of the Predator-Prey Model.

2 Preliminaries

In this section we will give a brief description of the ADM. For a detailed study, we refer the reader to [1]. The method has been applied to solve

many nonlinear equations and is usually characterized by its higher degree of accuracy using only few terms locally.
Consider the following type of nonlinear equation

$$L(\mathbf{u}) = R(\mathbf{u}) + N(\mathbf{u}), \tag{2.1}$$

where $L(\mathbf{u})$, $R(\mathbf{u})$ are linear operators and $N(\mathbf{u})$ is a nonlinear operator. Define $L^{-1}f = \int_0^t f(\tau)d\tau$ as the corresponding linear integrating operator of L.
The decomposition technique consists of representing the solution as an infinite series, namely,

$$u_1 = \sum_{n=0}^{\infty} u_{1n}, \qquad u_2 = \sum_{m=0}^{\infty} u_{2m}, \tag{2.2}$$

under the assumption of the convergence of solutions. Also the nonlinear operator N is decomposed as follows:

$$N(u_1) = \sum_{n=0}^{\infty} A_n, \qquad N(u_2) = \sum_{m=0}^{\infty} B_m, \tag{2.3}$$

where $A_n = A_n(u_{10}, u_{11}, ..., u_{1n}), B_m = B_m(u_{20}, u_{21}, ..., u_{2m})$ are the so-called *Adomian polynomials*. The convergence of ADM has been considered by Répaci and Cherruault locally (see [8] & [2]). We know that the u_{1n}s or u_{2m}s can be obtained in a recursive manner. Adomian polynomials are obtained by reordering and rearranging of the terms of the given equations according to the order which actually depends on both the subscripts and the powers of the u_{1n}s or u_{2m}s. In computing A_ns and B_ms, we note that a computer algebra system, Maple, comes in handy. It is the case in many nonlinear problems that one needs to compute few of these polynomials to obtain good approximate solutions locally.

3 Adapting The Adomian Decomposition Method (ADM) To Predator-Prey Model

Consider the coupled first order nonlinear differential equations (see [5] & [6]):

$$\frac{dQ(t)}{dt} = Q(t)\left\{r(1 - \frac{Q(t)}{K}) - \frac{kP(t)}{Q(t)+D}\right\},$$
$$\frac{dP(t)}{dt} = P(t)\left\{s(1 - \frac{hP(t)}{Q(t)})\right\}, \tag{3.1}$$

where r, K, k, D, s, h are positive constants subject to the initial conditions

$$Q(t_0) = Q_0, \qquad P(t_0) = P_0, \tag{3.2}$$

where $Q(t)$ and $P(t)$ are prey and predator populations at time t. Equations (3.1) can be rewritten in the non-dimensional form

$$\begin{aligned} \frac{du_1(\tau)}{d\tau} &= u_1(\tau)(1 - u_1(\tau)) - \frac{au_1(\tau)u_2(\tau)}{u_1(\tau) + d}, \\ \frac{du_2(\tau)}{d\tau} &= bu_2(\tau)(1 - \frac{u_2(\tau)}{u_1(\tau)}), \end{aligned} \tag{3.3}$$

with initial conditions

$$u_1(\tau_0) \equiv c_0 > 0, \qquad u_2(\tau_0) \equiv k_0 > 0, \tag{3.4}$$

when we nondimensionalize variables as (See [6] & [4]):

$$\begin{aligned} u_1(\tau) = \frac{Q(t)}{K}, \quad u_2(\tau) = \frac{hP(t)}{K}, \\ \tau = rt, \quad a = \frac{k}{hr}, \quad b = \frac{s}{r}, \quad d = \frac{D}{K}, \quad \tau_0 = rt_0. \end{aligned} \tag{3.5}$$

The existence and uniqueness of solutions of Equations (3.3) have been proved in [9]. Let

$$\begin{aligned} L(\mathbf{u}) &= \begin{pmatrix} \dfrac{du_1}{d\tau} \\ \dfrac{du_2}{d\tau} \end{pmatrix}, \\ R(\mathbf{u}) = \begin{pmatrix} u_1 \\ bu_2 \end{pmatrix}, \quad N(\mathbf{u}) &= \begin{pmatrix} -u_1^2 - \dfrac{au_1u_2}{u_1 + d} \\ -b\dfrac{u_2^2}{u_1} \end{pmatrix}. \end{aligned} \tag{3.6}$$

Integrating Equations (3.3) using L^{-1}, we get

$$\begin{aligned} u_1(\tau) - u_1(0) &= -\int_0^\tau u_1^2(v)dv - a\int_0^\tau \frac{u_2(v)u_1(v)}{d + u_1(v)}dv + \int_0^\tau u_1(v)dv, \\ u_2(\tau) - u_2(0) &= -b\int_0^\tau \frac{u_2^2(v)}{u_1(v)}dv + b\int_0^\tau u_2(v)dv. \end{aligned} \tag{3.7}$$

From Equation (3.3), if we apply the ADM we discussed in Sec 2 into Equations (3.7), then we get

$$u_1(\tau) = u_1(0) - \int_0^\tau u_1^2(v)dv - a\int_0^\tau \sum_{i=0}^{\infty} A_i dv + \int_0^\tau u_1(v)dv,$$
$$u_2(\tau) = u_2(0) - b\int_0^\tau \sum_{i=0}^{\infty} B_i dv + b\int_0^\tau u_2(v)dv, \qquad (3.8)$$

where

$$\begin{aligned}
&\textstyle\sum_{i=0}^{\infty} A_i = (u_{20}+u_{21}+u_{22}+...)(u_{10}+u_{11}+u_{12}+...)(p_0+p_1+p_2+...),\\
&\textstyle\sum_{i=0}^{\infty} B_i = (u_{20}+u_{21}+u_{22}+...)^2(q_0+q_1+q_2+...),
\end{aligned} \qquad (3.9)$$

with u_{1i} and u_{2i} $(i = 0, 1, 2, ...)$ defined in (2.2) and

$$p_0 = \frac{1}{d+u_{10}}, p_1 = -\frac{u_{11}}{(d+u_{10})^2}, p_2 = -\frac{u_{12}}{(d+u_{10})^2} + \frac{u_{11}^2}{(d+u_{10})^3},$$
$$p_3 = -\frac{u_{13}}{(d+u_{10})^2} + \frac{2u_{11}u_{12}}{(d+u_{10})^3} - \frac{u_{11}^3}{(d+u_{10})^4}, ...,$$
$$q_0 = \frac{1}{u_{10}}, q_1 = -\frac{u_{11}}{u_{10}^2}, q_2 = -\frac{u_{12}}{u_{10}^2} + \frac{u_{11}^2}{u_{10}^3},$$
$$q_3 = -\frac{u_{13}}{u_{10}^2} + \frac{2u_{11}u_{12}}{u_{10}^3} - \frac{u_{11}^3}{u_{10}^4}, ... \qquad (3.10)$$

Now

$$\begin{aligned}
&A_0 = u_{10}u_{20}p_0,\\
&A_1 = u_{20}u_{10}p_1 + u_{20}u_{11}p_0 + u_{21}u_{10}p_0,\\
&A_2 = u_{22}u_{10}p_0 + u_{21}u_{11}p_0 + u_{21}u_{10}p_1 + u_{20}u_{10}p_2 + u_{20}u_{12}p_0 + u_{20}u_{11}p_1,\\
&A_3 = u_{23}u_{10}p_0 + u_{22}u_{11}p_0 + u_{22}u_{10}p_1 + u_{21}u_{12}p_0 + u_{21}u_{11}p_1\\
&+u_{21}u_{10}p_2 + u_{20}u_{13}p_0 + u_{20}u_{12}p_1 + u_{20}u_{11}p_2 + u_{20}u_{10}p_3,\\
&........\\
&B_0 = q_0u_{20}^2,\\
&B_1 = 2u_{20}u_{21}q_0 + u_{20}^2q_1,\\
&B_2 = u_{21}^2q_0 + 2u_{20}u_{22}q_0 + 2u_{21}u_{20}q_1 + u_{20}^2q_2,\\
&B_3 = u_{20}^2q_3 + 2u_{20}u_{21}q_2 + 2u_{20}u_{22}q_1 + 2u_{20}u_{23}q_0 + 2u_{21}u_{22}q_0 + u_{21}^2q_1,\\
&.........
\end{aligned} \qquad (3.11)$$

Note that we compute the nonlinear term, $N(\mathbf{u})$, differently with [4] to improve the accuracy of computation. Let's choose the parameters of Equation (3.7) as $a = 1, b = 0.5, d = 0.01, c_0 = 0.6, k_0 = 0.2$ for a unique solution exhibiting limit cycle behavior (See [6]). Table 1 shows the comparison between the local approximate solutions obtained by the RK 4th-order method and the ADM with the initial 4 iterates.

τ	u_1(RK4)	u_1(ADM)	u_2(RK4)	u_2(ADM)
0.1	.6039553530	.6039553561	.2067331288	.2067331295
0.2	.6071656778	.6071657530	.2135943686	.2421461036
0.3	.6096326986	.6096332597	.2205756771	.2205758055
0.4	.6113602191	.6113625707	.2276678658	.2276684048
0.5	.6123538607	.6123610059	.2348605398	.2348621791

Table 1. Comparison between RK 4th-order Method and ADM when $a = 1, b = 0.5, d = 0.01, c_0 = 0.6, k_0 = 0.2$ (a limit cycle behavior).

We can see the absolute error is less than 10^{-5} for $\tau \in [0, 0.5]$ in both u_1 and u_2. To compare the accuracy of the ADM to the RK 4th-order method in an extended interval between $0 \leq \tau \leq 60$, we use a concatenation of the low-order ADM series solutions over a fixed subinterval with the length 0.5. To compute the approximate solutions of the model using both methods in the interval, we take totally 600 steps for both methods (see Table 2 for the comparison numerically). The graphic results of the ADM and the RK 4th-order method of u_1 and u_2 are in Figure 1.

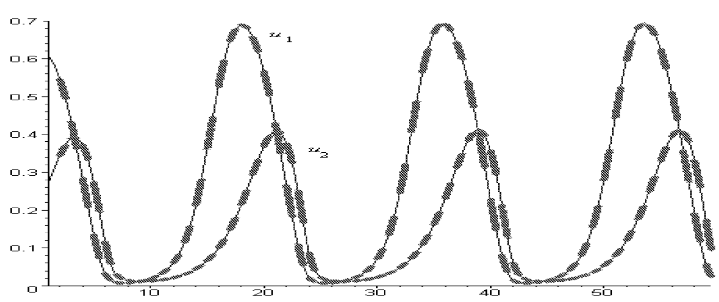


Figure 1. Comparison of the ADM solution and the RK 4th-order solution for an extended interval, $0 \leq \tau \leq 60$; (Solid line: solution of RK 4th-order Method; Dotted line: solution of ADM).

τ	u_1(RK4)	u_1(ADM)	u_2(RK4)	u_2(ADM)
10.0	.0185373	.0185369	.03	.0129918
20.0	.566519	.566516	.3683670	.368368
30.0	.0700825	.0700817	.0227728	.0227728
40.0	.228379	.228362	.356866	.356858
50.0	.278973	.27898	.0528530	.0528541
60.0	.0113245	.0106	.0438526	.0444

Table 2. Comparison of the RK 4th-order Method and ADM for an extended interval, $0 \leq \tau \leq 60$ when $a = 1, b = 0.5, d = 0.01, c_0 = 0.6, k_0 = 0.2$ (a limit cycle behavior).

Similarly, if we choose the parameters of Equation (3.7) as $a = 1, b = 5, d = 0.2, c_0 = 0.36, k_0 = 0.36$, the solution exhibits a stationary behavior (See [6]). Again, Table 3 shows the comparison between the local approximate solutions obtained by the RK 4th-order method and the ADM with the initial 4 iterates.

τ	u_1(RK4)	u_1(ADM)	u_2(RK4)	u_2(ADM)
0.1	.3598973665	.3598973665	.3599780636	.3599780642
0.2	.3597967288	.3597965724	.3599247372	.3599236983
0.3	.3596996508	.3596984918	.3598535146	.3598458009
0.4	.3596070021	.3596023794	.3597730419	.3597422521
0.5	.3595192089	.3595058717	.3596888052	.3595999147

Table 3. Comparison between the RK 4th-order Method and ADM when $a = 1, b = 5, d = 0.2, c_0 = 0.36, k_0 = 0.36$ (a stationary behavior).

We can see the absolute error is less than 10^{-4} for $\tau \in [0, 0.5]$ in both u_1 and u_2. In both cases, we compare the differences between the ADM solutions and the RK 4th-order solutions like Edwards et. al. ('99) and show much improved results for the ADM (See Table 4 & 5).

Therefore, ADM is still an effective method to provide reasonable approximate analytic solutions locally for the Predator-Prey Model depending on different computational methods for the nonlinear parts of the model equations.

	l_1 norm	l_2 norm	l_∞ norm
u_1	.009768493260	.001418880385	.00072451578
u_2	.01399205607	.1182892883	.00121071872

Table 4. Comparison of difference between the RK 4th-order Method and ADM in a limit cycle behavior.

	l_1 norm	l_2 norm	l_∞ norm
u_1	.0000083066	.000004801938	.0000046227
u_2	.0000503330	.007094575393	.0000307898

Table 5. Comparison of difference between RK 4th-order Method and ADM in a stationary behavior.

References

[1] G. Adomian, A Review of the Decomposition Method and Some Recent Results for Nonlinear Equations, *Computers Math. Applic.*, 21(5),101–127(1991).

[2] Y. Cherruault, Convergence of Adomian's Method, *Kybernetes*, 18(2), 31–38(1989).

[3] E. Deeba and J. Yoon, A Decomposition Method for Solving Nonlinear Oscillating Equations, *Int. Math. Journal*, 2(7), 813–823(2002).

[4] J. T. Edwards and J. A. Roberts and N. J. Ford, A Comparison of Adomian's Decomposition Method and Runge Kutta Methods for Approximate Solution of Some Predator Prey Model Equations, *International Journal of Applied Science and Computations*,2, 80-88(1999).

[5] R. M. May, Stability and Complexity in Model Ecosystems,*Princeton University Press*(2nd ed.), Princeton, N.J., 1974.

[6] J. D. Murray, *Mathematical Biology*, Springer-Verlag, 1987, 70–85.

[7] S. Olek, An Accurate Solution to the Multispecies Lotka-Volterra Equations, *SIAM Review*, 36(3), 480–488(1994).

[8] A. Répaci, A Nonlinear Dynamical Systems: On the Accuracy of Adomian's Decomposition Method, *Appl. Math Lett.*, 3(4), 35–39(1990).

[9] T. L. Saaty and J. Bram, Nonlinear Mathematics, *Dover*, 1964.

JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS,VOL.5,NO.4,331-336,2007,COPYRIGHT 2007 EUDOXUS PRESS,LLC

A FIXED POINT THEOREM FOR MULTIVALUED MAPPINS IN UNIFORM SPACE

D. TURKOGLU, H. ASLAN, AND S. N. MISHRA

ABSTRACT. In this paper we prove a new fixed point theorem for multi-valued mappings with an implicit relation on an orbitally complete uniform space.

1. INTRODUCTION

Uniform spaces form a natural extension of metric spaces. An exact analogue of the well-known Banach contraction principle in uniform spaces was obtained independently by Acharya [1] and Tarafdar [20]. Since then a number of fixed point theorems for single-valued and multi-valued mappings using various contractive conditions in this setting have been obtained ([2],[6], [9-12], [14-15], [17], [20-21], [23-24]). In this paper we first prove a fixed point theorem for a multi-valued mapping from an orbitally complete uniform space to its hyperspace. Subsequently, an application to locally convex spaces is also presented.

Let (X, u) be a uniform space. A family $P = \{d_i : i \in I\}$ of pseudometrics on X with indexing set I, is called an associated family for the uniformity u if the family

$$\beta = \{V(i, r) : i \in I, r > 0\}$$

where

$$V(i,r) = \{(x, y) : x, y \in X, d_i(x, y) < r\},$$

is a subbase for the uniformity u. We may assume β itself to be base by adjoining finite intersections of members of β, if necessary. The corresponding family of pseudometrics is called an augmented associated family for u. An augmented associated family for u will be denoted by P^*. For details the reader is referred to Tarafdar [20] and Thron [22]. Now onward, unless otherwise stated, X. will denote a uniform space (X, u) defined by P^*.

Let A be a nonempty subset of a uniform space X. Define

$$\Delta^*(A) = \sup\{d_i(x, y) : x, y \in A, i \in I\},$$

where $\{d_i(x, y) : i \in I\} = P^*$. Then $\Delta^*(A)$ is called an augmented diameter of A. Further, A is said to be P^*-bounded if $\Delta^*(A) < \infty$ (see [9]). Let

$$2^X = \{A : A \text{ is a nonempty } P^* - \text{bounded subset of } X\}.$$

2000 *Mathematics Subject Classification.* Primary 54H25; Secondary 47H10.

Key words and phrases. Fixed point, multi-valued mapping, implicit relation, uniform space, locally convex spaces.

For any nonempty subsets A and B of X, define

$$d_i(x, A) = \inf\{d_i(x, a) : a \in A, i \in I\}$$

$$\delta_i(A, B) = \sup\{d_i(a, b) : a \in A, b \in B, i \in I\}.$$

Let (X, u) be a uniform space and let $U \in u$ be an arbitrary entourage. Then the uniformity 2^u on 2^X is defined by the base

$$2^\beta = \{\tilde{U} : U \in u\}$$

where

$$\tilde{U} = \{(A, B) \in 2^X \times 2^X : A \times B \subseteq U\} \cup \Delta$$

(Here Δ denotes the diagonal of $X \times X$).

The augmented associated family P^* also induces a uniformity u^* on 2^X defined by the base

$$\beta^* = \{V^*(i, r) : i \in I, r > 0\},$$

where

$$V^*(i, r) = \{(A, B) \in 2^X \times 2^X : \delta_i(A, B) < \varepsilon\} \cup \Delta.$$

The uniformities 2^u and u^*on 2^X are uniformly isomorphic. The space $(2^X, u^*)$ is thus a uniform space called the hyperspace of (X, u). There exist other bases which could be used to generate uniformities on X as well as on 2^X (see, for details [3], [14], [15]).

We recall the following definitions and notations from [4], [5] and [19].

A sequence$\{A_n\}$ of sets in 2^X is said to converge to the subset A of X if the following two conditions are satisfied:

(i) For each point in a in A, there is a sequence $\{a_n\} \subseteq A_n$ for all n and $a_n \to a$.

(ii) For every $\varepsilon > 0$, there is an integer N such that $A_n \subseteq A_\varepsilon$ for $n \geq N$, where

$$A_\varepsilon = \underset{x \in A}{\cup} U(x) = \{y \in X : d_i(x, y) < \varepsilon \text{ for some } x \text{ in } A,\ i \in I\}.$$

In such a case, A is said to be limit of the sequence $\{A_n\}$ and we write $\lim\limits_{n\to\infty} A_n = A$ or $A_n \to A$ as $n \to \infty$.

(iii) A multi-valued mapping $F : X \to 2^X$ is said to be continuous at $x_0 \in X$ if whenever $\{x_n\}$ is a sequence of points in X converging to x, the sequence $\{Fx_n\}$ in 2^X converges to Fx in 2^X. We say that F is a continuous mapping of X into 2^X if F is continuous at each point x in X.

(iv) An orbit of a multi-valued mapping F :$X \to 2^X$ at a point $x_0 \in X$ is a sequence $\{x_n\}$ given by

$$O(F, x_0) = \{x_n : x_n \in Fx_{n-1}, n = 1, 2, 3, ...\}.$$

A uniform space X is said to be $F-$orbitally complete if every Cauchy sequence which is a subsequence of an orbit of F at each $x \in X$ converges to a point of X. The above notions of an orbit and orbital completeness were introduced first by Smithson [19] and Ciric [4] respectively in the context of a metric space. There exist orbitally complete spaces which are not complete (see [4]).

(v) $F : X \to 2^X$ is called point closed if Fx is a closed subset of X for each $x \in X$.

For fixed point theory of multi-valued mappings, we refer to Hicks [7] and Rhoades [8] and references thereof (also, see [13], [16] for some related results)

Theorem 1. [13] *Let F be a continuous mapping of an $F-$orbitally complete metric space (X, d) into 2^X satisfying*

$$\min\{\delta(\overline{F}x, \overline{F}y)^r, \delta(x, \overline{F}x)\delta(y, \overline{F}y)^{r-1}, \delta(y, \overline{F}y)^r\} + a\min\{d(x, \overline{F}y), d(y, \overline{F}x)\} \leq (pd(x, \overline{F}x) + qd(x, y))d(y, \overline{F}y)^{r-1}$$

for all x, y in X, where $r \geq 1$ is an integer, a, p, q are real numbers such that $0 < p + q < 1$. Then there exists $x \in X$ such that $x \in \overline{F}x$ (the closure of Fx). If F is a point closed mapping, then $x \in Fx$, that is, x is a fixed point of F.

2. Implicit Relations

Let $\Im_8$ be the set of all functions $\Phi : (0, \infty)^8 \to \mathbb{R}$ satisfying the following conditions:

1^0. Φ is nondecreasing in variables t_1, t_3, t_4 and decreasing in variables t_5, t_6, t_7;
2^0. There exists $h \in [0, 1)$ such that $\Phi(u, v, v, u, v, u, u+v, 0) \leq 0$ implies $u \leq hv$.

Example 1. *Let*

$$\Phi(t_1, ...t_8) = \min\{t_1^r, t_3t_4^{r-1}, t_4^r\} - a\min\{t_7, t_8\} - (pt_5 + qt_2)t_6^{r-1},$$

where $r \in \mathbb{N}^*$(the non-negative integers), $a > 0, 0 < p + q < 1$. Then
1^0. Obviously.
2^0. $\Phi(u, v, v, u, v, u, u + v, 0) \leq 0$ implies $\min\{u^r, u^{r-1}v\} - (p + q)u^{r-1}v \leq 0$. If $u \geq v$ then $u^{r-1}v(1 - p - q) \leq 0$ which implies $p + q \geq 1$ a contradiction. Thus $u < v$ and then $u^r - (p + q)\ u^{r-1}v \leq 0$ which implies $u \leq hv$, where $h = (p + q) \in (0, 1)$.

3. Main Results

Theorem 2. *Let F be a continuous mapping of an $F-$orbitally complete Hausdorff uniform space (X, u) into 2^X satisfying*

$$\Phi(\delta_i(\overline{F}x, \overline{F}y), d_i(x, y), \delta_i(x, \overline{Fx}), \delta_i(y, \overline{F}y), d_i(x, \overline{F}x),$$

$$d_i(y, \overline{Fy}), d_i(x, \overline{F}y), d_i(y, \overline{Fx})) \leq 0 \tag{3.1}$$

for all $x, y \in X, i \in I$ where $\Phi \in \Im_8$ then there exists $x \in X$ such that $x \in \overline{F}x$. If F is a point closed mapping, then $x \in Fx$ that is, x is a fixed point of F.

Proof. Let x_0 be an arbitrary point in X and consider the sequence $\{x_n\}$ defined by

$$x_1 \in \overline{F}x_0, x_2 \in \overline{F}x_1, ..., x_n \in \overline{F}x_{n-1}, n = 1, 2, ...$$

Let us suppose that $d_i(x_n, \overline{F}x_n) > 0$ for each $n = 0, 1, 2, ...$ (otherwise for some positive integer, $x_n \in \overline{F}x_n$ as desired) such that $d_i(x_n, x_{n+1}) > 0, n = 0, 1, 2, ...$. For $x = x_{n-1}$ and $y = x_n$ by condition (3.1) we have

$$\Phi(\delta_i(\overline{F}x_{n-1}, \overline{F}x_n), d_i(x_{n-1}, x_n), \delta_i(x_{n-1}, \overline{F}x_{n-1}), \delta_i(x_n, \overline{F}x_n),$$

$$d_i(x_{n-1}, \overline{F}x_{n-1}), d_i(x_n, \overline{F}x_n), d_i(x_{n-1}, \overline{F}x_n), d_i(x_n, \overline{F}x_{n-1})) \leq 0.$$

Hence from the definition of δ_i and property 1^0

$$\Phi(d_i(x_n,x_{n+1}),d_i(x_{n-1},x_n),d_i(x_{n-1},x_n),d_i(x_n,x_{n+1}),$$
$$d_i(x_{n-1},x_n),d_i(x_n,x_{n+1}),d_i(x_{n-1},x_{n+1}),0)\le 0,$$
$$\Phi(d_i(x_n,x_{n+1}),d_i(x_{n-1},x_n),d_i(x_{n-1},x_n),d_i(x_n,x_{n+1}),$$
$$d_i(x_{n-1},x_n),d_i(x_n,x_{n+1}),d_i(x_{n-1},x_n)+d_i(x_n,x_{n+1}),0)\le 0$$

which implies by 2^0 that

$$d_i(x_n,x_{n+1})\le hd_i(x_{n-1},x_n), n=1,2,\ldots.$$

Now it can be easily verified that $\{x_n\}$ is a Cauchy sequence. Since (X,u) is $F-$orbitally complete, there exists a point $x\in X$ such that $x_n\to x$. By the continuity of F, the sequences $\{Fx_n\}$ in 2^X converge to Fx in 2^X.

Next we shall prove that $x\in\overline{F}x$. Let $W\in u$ be arbitrary and let $V(j,t)\in\beta$, $j\in I$ and $t>0$ be such that $V(j,t)\subseteq W$. For a given $\varepsilon>0$, there exists a positive integer N_1 such that

$$d_j(x_n,x)<\frac{\varepsilon}{3} \tag{3.2}$$

for all $n\ge N_1$. On the other hand, since F is continuous, for the same ε, we can find an N_2 such that

$$Fx_n\subseteq A_{\frac{\varepsilon}{3}}=\underset{a\in Fx}{\cup}U(a)$$

for all $n\ge N_2$. Further, since $x_{n+1}\in\overline{F}x_n$, then there exists a $y\in Fx_n$ such that $d_j(x_n,y)<\frac{\varepsilon}{3}$, and

$$y\in\overline{F}x_n\subseteq\underset{a\in Fx}{\cup}U(a)$$

implies that there exists an $a\in Fx$ such that $d_j(a,y)<\frac{\varepsilon}{3}$. Thus

$$d_j(x_{n+1},Fx)\le d_j(x_n,a)\le d_j(x_n,y)+d_j(y,a)<\frac{\varepsilon}{3}+\frac{\varepsilon}{3}=\frac{2\varepsilon}{3} \tag{3.3}$$

for all $n\ge N_2$. Now let $N=\max\{N_1,N_2\}$, from (3.2) and (3.3), we have

$$d_j(x,Fx)\le d_j(x,x_{n+1})+d_j(x_{n+1},Fx)<\frac{\varepsilon}{3}+\frac{2\varepsilon}{3}=\varepsilon$$

for all $n\ge N$. Since $\varepsilon>0$ is arbitrary, we have $d_j(x,Fx)=0<t$. Hence $(x,Fx)\in V(j,t)\subseteq W$. It follows that $x\in\overline{F}x=Fx$, since Fx is closed for each $x\in X$. Hence x is a fixed point of F. This completes the proof of the theorem. □

Corollary 1. *Let F be a continuous mapping of an F-orbitally complete Hausdorff uniform space (X,u) into 2^X satisfying*

$$\min\{\delta_i(\overline{F}x,\overline{F}y)^r,\delta_i(x,\overline{F}x)\delta_i(y,\overline{F}y)^{r-1},\delta_i(y,\overline{F}y)^r\}+a\min\{d_i(x,\overline{F}y),d_i(y,\overline{F}x)\}$$
$$\le(pd_i(x,\overline{F}x)+qd_i(x,y))d_i(y,\overline{F}y)^{r-1}$$

for all x,y in $X,i\in I$ where $r\ge 1$ is an integer, a,p,q are real numbers such that $a>0$ and $0<p+q<1$, then there exists $x\in X$ such that $x\in\overline{F}x$. If F is a point closed mapping, then F has fixed point.

Proof. It follows from Theorem 2 and Example 1. □

Corollary 2. *Theorem 1 for $a>0$.*

Proof. It follows as a special of Corollary 1 by replacing the uniform space (X,u) by a metric space (X,d). □

4. Application to locally convex spaces

Let (X, τ) be a locally convex linear topological space whose topology is τ generated by a family of seminorms $\{p_i : i \in I\}$ so that the collection

$$\{V(i,r) : \ i \in I, \ r > 0\},$$

where $V(i,r) = \{x \in X : p_i(x) < r\}$ is a neighborhood base for τ. Then the family $P^* = \{ \ p_i : i \in I\}$ is called an augmented associated family for τ.

Now, for each $i \in I$, the function $d_i : X \times X \to \mathbb{R}$ defined by $d_i(x,y) = p_i(x-y)$ for all $x, y \in X$ is a pseudometric on X. Thus the family $P^* = \{p_i : i \in I\}$ determines a unique uniformity u on X and the uniform topology of X coincides with the locally convex topology τ of the space (see Shaefer [18]).

For any nonempty subsets A and B of X, we have

$$d_i(x, A) = \inf\{p_i(x-a) : a \in A, i \in I\},$$

$$\delta_i(A, B) = \sup\{p_i(a-b) : a \in A, b \in B, i \in I\}. \tag{4.1}$$

Then using an idea of Tarafdar [21] we have the following result as an application of Theorem 2.

Theorem 3. *Let F be a point closed continuous mapping of an F-orbitally complete Hausdorff locally convex linear topological space (X, τ) into 2^X satisfying the conditions of Theorem 2 with d_i and δ_i as indicated above (4.1). Then F has a fixed point in X.*

Remark 1. *Analogues of Corollaries 1 and 2 can also be formulated easily.*

References

[1] S. P. Acharya, Some results on fixed points in uniform space, Yokohama. Math. J. 22(1974), 105-116.
[2] V.V. Angelov, Fixed point theorem in uniform spaces and applications, Czechoslovak Math. J. 37(112)(1997), 19-32.
[3] Titu Banzaru and Bela Rendi, Topologies on spaces of subsets and multivalued mappings, Mathematical Monographs 63, University of Timisora, 1997.
[4] L. B. Ciric, On contractive type mappings, Math. Balkanica. 1(1971), 52-57.
[5] B. Fisher, Common fixed points of mappings and multivalued mappings, Rostock Math. Kolloq. 18(1981), 69-77.
[6] A. Ganguly, Fixed point theorems for three maps in uniform spaces, Indian J. Pure Appl. Math. 17(4)(1986), 476-480.
[7] T. L. Hicks, Set Valued Mappings on Metric Space, Indian J. Pure Appl. Math. 22(4)(1999), 269-271.
[8] T. L. Hicks and B. E. Rhoades, Fixed points and continuity for multivalued mappings, Internat. J. Math. Math. Sci. 15(1)(1992), 15-30.
[9] S. N. Mishra, A note on common fixed points of multivalued mappings in uniform spaces, Math. Sem. Notes Kobe Univ. 9(1981), 341-347.
[10] S. N. Mishra, On common fixed points of multimappings in uniform spaces, Indian J. Pure Appl. Math.13(5)(1982), 606-608.
[11] S. N. Mishra, Fixed points of contractive type multivalued mappings in uniform spaces, Indian J. Pure Appl. Math. 18(4)(1987), 283-289.
[12] S. N. Mishra and S. L. Singh, Fixed points of multivalued mappings in uniform spaces, Bull. Cal. Math. Soc. 77(1985), 323-329.
[13] O. Özer and , D. Türkoğlu, Some fixed point theorems for multivalued mappings, Anadolu Üniversitesi, Fen Fakültesi Dergisi (1998), Sayı 4, 83-96.

[14] D. V. Pai and P. Veeramani, Fixed point theorems for multi-mappings, Yokohama Math. J. 28(1980), 7-14.
[15] D. V. Pai and P. Veeramani, Fixed point theorems for multi-mappings, Indian J. Pure Appl. Math. 11(7)(1980), 891-896.
[16] V. Popa and D. Türkoğlu, Some Fixed Point Theorems For Hybrid Contractions Satisfying an Implicit Relation, Universitatea Din Bacău Studii Şi Cercetări Ştiinţifice Seria: Matematică Nr. 8 (1998) pag. 75-86.
[17] K. Qureshi and S. Upadhyay, Fixed point theorems in uniform spaces, Bull. Calc. Math. Soc., 84(1992), 5-10.
[18] H. H. Shaefer, Topological vector spaces, Macmillan, NewYork, 1966.
[19] R. E. Smithson, Fixed points for contractive multi-functions, Proc. Amer. Math. Soc. 27(1971), 192-194.
[20] E. Tarafdar, An approach to fixed point theorems on uniform spaces, Trans. Amer. Math. Soc. 191(1974), 209-225.
[21] E. Tarafdar, On a fixed point theorem on locally convex linear topological spaces, Monatshefte für Mathematik 82(1976), 341-344.
[22] W. J. Thron, Topological structures, Holt, Rinehart and Winston, New York, 1966.
[23] D. Türkoğlu, O. Özer and B. Fisher, Some fixed point theorems for set valued mappings in uniform spaces, Demonstratio Math, XXXII, 2, (1999), 395-400.
[24] C. S. Wong, A fixed point theorem for a class mappings, Math. Ann. 204(1973), 97-10.

Department of Mathematics, Faculty of Science and Arts, University of Gazi, 06500-Teknikokullar Ankara, TURKEY
E-mail address: dturkoglu@gazi.edu.tr

Department of Mathematics, Faculty of Science and Arts, University of Gazi, 06500-Teknikokullar Ankara, TURKEY
E-mail address: haslan@gazi.edu.tr

Department of Mathematics, University of Transkei Umtata 5100, SOUTH AFRICA
E-mail address: mishra@getafix.utr.ac.za

JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS,VOL.5,NO.4,337-346,2007,COPYRIGHT 2007 EUDOXUS PRESS,LLC

SOME RESULTS ON STABILITY FOR PERTURBED LINEAR DYNAMIC SYSTEMS ON TIME SCALES.

İ.B.YAŞAR, A.TUNA AND M.T. DASTJERDI

Abstract. we study some results on stability for perturbed linear systems by using Unified Gronwall's inequality on time scales.

1. Introduction

The theory of dynamic equations on time scales (aka measure chains) was introduced by Hilger [6] with the motivation of providing a unified approach to continuous and discrete analysis. The generalized derivative or Hilger derivative $f^{\triangle}(t)$ of a function $f : \mathbb{T} \to \mathbb{R}$, where $\mathbb{T}$ is a so-called "time scale" (an arbitrary closed nonempty subset of $\mathbb{R}$) becomes the usual derivative when $\mathbb{T} = \mathbb{R}$, that is $f^{\triangle}(t) = f'(t)$. On the other hand, if $\mathbb{T} = \mathbb{Z}$, then $f^{\triangle}(t)$ reduces to the usual forward difference, that is $f^{\triangle}(t) = \triangle f(t)$. This theory not only brought equations leading to new applications. Also, this theory allows one to get some insight into and better understanding of the subtle differences between discrete and continuous systems [1, 3].

Dacunha [4] have introduced stability for time varying linear dynamic systems on time scales. He introduced the unified theorems of uniform stability and uniform exponential stability of linear systems on time scales, as well as illustrations of these theorems in examples, and demonstrated how the quadratic Lyapunov function developed, it can also be used to determine instability of a system.

In this paper, we study some results on stability for perturbed linear dynamic systems by using Unified Gronwall's inequality on time scales.

Now, first we mention without proof several foundational definitions and result from the calculus on time scales in an excellent introductory text by Bohner and Peterson [3].

2. General Definitions

A *time scale* $\mathbb{T}$ is any nonempty closed subset of the real numbers $\mathbb{R}$. Thus time scales can be any of the usual integer subsets (e.g. $\mathbb{Z}$ or $\mathbb{N}$), the entire real line $\mathbb{R}$, or any combination of discrete points unioned with continuous intervals. The majority of research on time scales so far has focused on expanding and generalizing the vast suite of tools available to the differential and difference equation theorist. We now briefly outline the portions of the time scales theory that are needed for this paper to be as self-contained as is practically possible.

2000 *Mathematics Subject Classification.* Primary 05C38, 15A15; Secondary 05A15, 15A18.

Key words and phrases. Stability; Time scale; Perturbed Linear system; Unified Gronwall's inequality.

The *forward jump operator* of $\mathbb{T}$, $\sigma(t) : \mathbb{T} \to \mathbb{T}$, is given by $\sigma(t) = \inf_{s\in\mathbb{T}} \{s > t\}$. The *backward jump operator* of $\mathbb{T}$, $\rho(t) : \mathbb{T} \to \mathbb{T}$, is given by $\rho(t) = \inf_{s\in\mathbb{T}} \{s < t\}$. The *graininess function* $\mu(t) : \mathbb{T} \to [0,\infty)$ is given by $\mu(t) = \sigma(t) - t$. Here we adopt the conventions $\inf \emptyset = \sup \mathbb{T}$ (i.e. $\sigma(t) = t$ if $\mathbb{T}$ has a maximum element t), and $\sup \emptyset = \inf \mathbb{T}$ (i.e. $\rho(t) = t$ if $\mathbb{T}$ has a minimum element t). For notational purposes, the intersection of a real interval $[a, b]$ with a time scale $\mathbb{T}$ is denoted by $[a, b] \cap \mathbb{T}$: $[a, b]_{\mathbb{T}}$.

A point $t \in \mathbb{T}$ is *right-scattered* if $\sigma(t) > t$ and *right dense* if $\sigma(t) = t$. A point $t \in \mathbb{T}$ is *left-scattered* if $\rho(t) < t$ and *left dense* if $\rho(t) = t$ If t is both left-scattered and right-scattered, we say t is *isolated.* If t is both left-dense and right-dense, we say t is dense. The set $\mathbb{T}^K$ is defined as follows: If $\mathbb{T}$ has a left-scattered maximum m, then $\mathbb{T}^K = \mathbb{T}$ -$\{m\}$; otherwise, $\mathbb{T}^K = \mathbb{T}$. If $f : \mathbb{T} \to\mathbb{R}$ is a function, then the composition $f(\sigma(t))$ is often denoted by $f^\sigma(t)$.

For $f : \mathbb{T} \to\mathbb{R}$ and $t \in \mathbb{T}^K$, define $f^\triangle(t)$ as the number (when it exists), with the property that, for any $\varepsilon > 0$, there exists a neighborhood U of f such that

$$\left|[f(\sigma(t)) - f(s))] - f^\triangle(t)\,[\sigma(t) - s]\right| \le \varepsilon\,|\sigma(t) - s|\,, \qquad \forall s \in U.$$

The $f^\triangle : \mathbb{T}^K \to \mathbb{R}$ is called the *delta derivative* or the *Hilger derivative* of f on $\mathbb{T}^K$. We say f is *delta differentiable* on $\mathbb{T}^K$ provided $f^\triangle(t)$ exists for all $t \in \mathbb{T}^K$.

The following theorem establishes several important observations regarding delta derivatives.

Theorem 1. *Suppose $f : \mathbb{T} \to\mathbb{R}$ and $t \in \mathbb{T}^K$.*

(i) If f is differentiable at t, then f is continuous at t.

(ii) If f is continuous at t and t is right-scattered, then f is delta differentiable at t and $f^\triangle(t) = \frac{f^\sigma(t)-f(t)}{\mu(t)}$.

(iii) If t is right-dense, then f is delta differentiable at t if and only if $\lim_{t\to s} \frac{f(t)-f(s)}{t-s}$ exists. In this case, $f^\triangle(t) = \lim_{t\to s} \frac{f(t)-f(s)}{t-s}$.

(iv) If f is delta differentiable at t, then $f(\sigma(t) = f(t) + \mu(t) f^\triangle(t)$.

Note that $f^\triangle$ is precisely f' from the usual calculus when $\mathbb{T} =\mathbb{R}$. On the other hand, $f^\triangle = \triangle f = f(t + 1) - f(t)$ (i.e. the forward difference operator) on the time scale $\mathbb{T} = \mathbb{Z}$. These are but two very special (and rather simple) examples of time scales. Moreover, the realms of differential equations and difference equations can now be viewed as but special, cases of more general *dynamic equations on time scales*, i.e. equations involving the delta derivative(s) of some unknown function.

A function $f : \mathbb{T} \to\mathbb{R}$ is *rd-continuous* if f is continuous at every right dense point $t \in \mathbb{T}$, and its left hand limit exists at each left dense point $t \in \mathbb{T}$. The set of rd-continuous functions $f : \mathbb{T} \to\mathbb{R}$ will be denoted by $C_{rd} = C_{rd}(\mathbb{T}) = C_{rd}(\mathbb{T},\mathbb{R})$. A function $F : \mathbb{T} \to\mathbb{R}$ is called a (delta) *antiderivative* of $f : \mathbb{T} \to\mathbb{R}$ provided $F^\triangle(t) = f(t)$ holds for all $t \in \mathbb{T}^K$. The *cauchy integral* or *definite integral* is given by $\int\limits_a^b f(t)\triangle t = F(b) - F(a)$, for all $a, b \in \mathbb{T}$, where F is any (delta) antiderivative of f. Suppose that $\sup \mathbb{T} = \infty$. Then the *improper integral* is defined to by $\int\limits_a^\infty f(t)\triangle t = \lim_{b\to\infty} F(t)\mid_a^b$ for all $a \in \mathbb{T}$. We remark that the integral is defined in terms of a Lebesgue type integral [2] or a Riemann integral [3].

Theorem 2. *(Existence of antiderivatives).*

(i) Every rd-continuous function has an antiderivative. If $t_0 \in \mathbb{T}$*, then* $F(t) = \int\limits_{t_0}^{t} f(\tau) \triangle \tau$, $t \in \mathbb{T}$*, is an antiderivate of* f*.*

(ii) If $f \in C_{rd}$ *and* $t \in \mathbb{T}^K$, $\int\limits_{t}^{\sigma(t)} f(\tau) \triangle \tau = f(\tau)\mu(\tau)$.

(iii) Suppose $a, b \in \mathbb{T}$ *and* $f \in C_{rd}$.

(a) If $\mathbb{T} = \mathbb{R}$*, then* $\int\limits_{a}^{b} f(t) \triangle t = \int\limits_{a}^{b} f(t)dt$ *(the usual Riemann integral).*

(b) If $[a, b]_{\mathbb{T}}$ *consists of only isolated points, then*

$$\int\limits_{a}^{b} f(t) \triangle t = \begin{cases} \sum\limits_{t\in[a,b)_{\mathbb{T}}} f(t)\mu(t), & a < b, \\ 0, & a = b, \\ -\sum\limits_{t\in[a,b)_{\mathbb{T}}} f(t)\mu(t) & a > b. \end{cases}$$

The last result above reveals that in the continuous case, $\mathbb{T} = \mathbb{R}$, definite integrals are usual Riemann integrals from calculus. When $\mathbb{T} = \mathbb{Z}$, definite integrals correspond to definite sums from the difference calculus ; see [9].

Theorem 3. *Let* $f : \mathbb{R} \to \mathbb{R}$ *be continuously differentiable and suppose that* $g : \mathbb{T} \to \mathbb{R}$ *is delta differentiable. Then* $f \circ g : \mathbb{T} \to \mathbb{R}$ *is delta differentiable and the formula*

$$(f \circ g)^{\triangle} (t) = g^{\triangle}(t) \int\limits_{0}^{1} f'(g(t) + \delta\mu(t)g^{\triangle}(t))d\delta$$

holds.

2.1. **The Hilger complex plane.** For $h > 0$, define the *Hilger complex numbers*, the *Hilger real axis*, the *Hilger alternating axis*, and the *Hilger imaginary circle* by

$\mathbb{C}_h = \left\{z \in \mathbb{C} : z \neq -\frac{1}{h}\right\}$, $\mathbb{R}_h = \left\{z \in \mathbb{R} : z > -\frac{1}{h}\right\}$

$\mathbb{A}_h = \left\{z \in \mathbb{R} : z < -\frac{1}{h}\right\}$, $\mathbb{I}_h = \left\{z \in \mathbb{C} : \left|z + \frac{1}{h}\right| = \frac{1}{h}\right\}$

respectively. For $h = 0$, let $\mathbb{C}_0 := \mathbb{C}$, $\mathbb{R}_0 := \mathbb{R}$, $\mathbb{A}_0 := \emptyset$, and $\mathbb{I}_0 := i\mathbb{R}$.

Let $h > 0$ and $z \in \mathbb{C}_h$. The *Hilger real part of* z is defined by $\mathrm{Re}_h(z) := \frac{|zh+1|}{h}$, and the *Hilger imaginary part of* z is defined by $\mathrm{Im}_h(z) := \frac{Arg(xh+1)}{h}$, where $Arg(z)$ denotes the principle argument of z (i.e.,$-\pi < Argz \leq \pi$).

For $h > 0$, define the strip $\mathbb{Z}_h := \left\{z \in \mathbb{C} : \frac{-\pi}{h} < Argz \leq \frac{\pi}{h}\right\}$, and for $h = 0$, set $\mathbb{Z}_0 := \mathbb{C}$. Then we can define the *cylinder transformation* $\xi_h = \mathbb{C}_h \to \mathbb{Z}_h$ by

$$\xi_h(z) = \frac{1}{h} Log(1 + zh), \quad h > 0$$

where Log is the principle logarithm function. When $h = 0$, we define $\xi_0(z) = z$,

for all $z \in \mathbb{C}$. It then follows that the *inverse cylinder transformation* $\xi_h^{-1} : \mathbb{Z}_h \to \mathbb{C}_h$ is given by

$$\xi_h^{-1}(z) = \frac{e^{zh} - 1}{h}.$$

Since the graininess may not be constant for a given time scale, we will interchangeably subscript various quantities (such as ξ and ξ^{-1}) with $\mu = \mu(t)$ instead of h to reflect this.

2.2. **Generalized exponential Functions.** The function $p : \mathbb{T} \to \mathbb{R}$ is *regressive* if $1 + \mu(t)p(t) \neq 0$ for all $t \in \mathbb{T}^K$, and this concept motivates the definition of the following sets:

$$\Re = \left\{p : \mathbb{T} \to \mathbb{R} : p \in C_{rd}(\mathbb{T}) \text{ and } 1 + \mu(t)p(t) \neq 0 \ \forall t \in \mathbb{T}^K\right\},$$

$$\Re^+ = \left\{p \in \Re : 1 + \mu(t)p(t) > 0 \text{ for all } t \in \mathbb{T}^K\right\}.$$

The function $p : \mathbb{T} \to \mathbb{R}$ is *uniformly regressive* on $\mathbb{T}$ there exists a positive constant δ such that $0 < \delta^{-1} \leq |1 + \mu(t)p(t)|$, $t \in \mathbb{T}^K$. A matrix is regressive if and only if all of its eigenvalues are in $\Re$. Equivalently, the matrix $A(t)$ is regressive if and only if $I + \mu(t)A(t)$ is invertible for all $t \in \mathbb{T}^K$.

If $p \in \Re$, then we define the *generalized time scale exponential function* by

$$e_p(t,s) = \exp\left(\int_s^t \xi_{\mu(\tau)}(p(\tau)) \triangle \tau\right) \quad \text{for all } s,t \in \mathbb{T}$$

The following theorem is a compilation of properties of $e_p(t,s)$ (some of which are counterintuitive) that we need in the main body of the paper.

Theorem 4. *The function $e_p(t,s)$ has the following properties:*

(i) If $p \in \Re$, then $e_p(t,r)e_p(r,s) = e_p(t,s)$ for all $r,s,t \in \mathbb{T}$.

(ii) $e_p(\sigma(t),s) = (1 + \mu(t)p(t))e_p(t,s)$.

(iii) If $p \in \Re^+$, then $e_p(t,t_0) > 0$ for all $t \in \mathbb{T}$.

(iv) If $1 + \mu(t)p(t) < 0$ for some $t \in \mathbb{T}^K$, then $e_p(t,t_0)e_p(\sigma(t),t_0) < 0$.

(v) If $\mathbb{T} = \mathbb{R}$, then $e_p(t,s) = e^{\int_s^t p(\tau)d\tau}$. Morever, If p is constant, then $e_p(t,s) = e^{p(t-s)}$.

(vi) If $\mathbb{T} = \mathbb{Z}$, then $e_p(t,s) = \Pi_{\tau=s}^{t-1}(1 + p(\tau))$. Morever, If $\mathbb{T} = h\mathbb{Z}$, with $h > 0$ and p is constant, then $e_p(t,s) = (1 + hp)^{\frac{(t-s)}{h}}$.

If $p \in \Re$ and $f : \mathbb{T} \to \mathbb{R}$ is rd-continuous, then the dynamic equation

$$y^{\triangle}(t) = p(t)y(t) + f(t) \tag{2.1}$$

is called regressive.

Theorem 5. *(Variation of constants). Let $t_0 \in \mathbb{T}$ and $y(t_0) = y_0 \in \mathbb{R}$. Then the regressive IVP (2.1) has a unique solution $y : \mathbb{T} \to \mathbb{R}^n$ given by*

$$y(t) = y_0 e_p(t,t_0) + \int_{t_0}^t e_p(t,\sigma(\tau))f(\tau) \triangle \tau.$$

We say the $n \times 1$-vector-valued system

$$y^{\triangle}(t) = p(t)y(t) + f(t) \tag{2.2}$$

is *regressive* provided $A \in \Re$ and $f : \mathbb{T} \to \mathbb{R}^n$ is rd-continuous vector-valued function.

Let $t_0 \in \mathbb{T}$ and assume that $A \in \Re$ is an $n \times n$-matrix-valued function. The unique matrix-valued solution to the IVP

$$Y^{\triangle}(t) = p(t)Y(t), \quad Y(t_0) = I_n \tag{2.3}$$

where I_n is the $n \times n$-identitiy matrix, is called the *transition matrix* and it is denoted by $\Phi_A(t, t_0)$.

In this paper, we denote the solution to (2.3) as $\Phi_A(t, t_0)$ when $A(t)$ is time varying and denote the solution as $e_A(t, t_0) \equiv \Phi_A(t, t_0)$ (the *matrix exponential*, as in [3]) only when $A(t) \equiv A$ is a constant matrix. Also, if $A(t)$ is a function on $\mathbb{T}$ and the time scale matrix exponential function is a function on some other time scale $\mathbb{S}$, then $A(t)$ is constant with respect to $e_{A(t)}(\tau, s)$, for all $\tau, s \in \mathbb{S}$ and $t \in \mathbb{T}$. The following lemma lists some properties of the transition matrix.

Theorem 6. *Suppose $A, B \in \Re$ are matrix-valued functions on $\mathbb{T}$.*

(i) Then the semigroup property $\Phi_A(t, r)\Phi_A(r, s) = \Phi_A(t, s)$ is satisfies for all $r, s, t \in \mathbb{T}$.

(ii) $\Phi_A(\sigma(t), s) = (1 + \mu(t)p(t))\Phi_A(t, s)$.

(iii) If $\mathbb{T} = \mathbb{R}$ and A is constant, then $\Phi_A(t, s) = e_A(t, s) = e^{A(t-s)}$.

(iv) If $\mathbb{T} = h\mathbb{Z}$, with $h > 0$ and A is constant, then $\Phi_A(t, s) = e_A(t, s) = (1 + hp)^{\frac{(t-s)}{h}}$.

We now present a theorem that guarantees a unique solution to the regressive $n \times 1$-vector-valued dynamic IVP (2.2) .

Theorem 7. *(Variation of constants). Let $t_0 \in \mathbb{T}$ and $y(t_0) = y_0 \in \mathbb{R}^n$. Then the regressive IVP (2.2) has a unique solution $y : \mathbb{T} \to \mathbb{R}^n$ given by*

$$y(t) = y_0\Phi_A(t, t_0) + \int_{t_0}^{t} \Phi_A(t, \sigma(\tau))f(\tau) \triangle \tau. \tag{2.4}$$

Beside the calculus on the time scales, we talk about several foundational definitions and result from the stability for time varying linear dynamic systems on time scales in an excellent introductory text by DaCunha [4].

3. Stability

We start introducing definitions and notation that will be employed in the sequel. The *Euclidean norm* of an $n \times 1$ vector $x(t)$ is defined to be a real-valued function of t and is denoted by

$$\|x(t)\| = \sqrt{x^T(t)x(t)}.$$

The induced norm of an $m \times n$ matrix A is defined to be

$$\|A(t)\| = \max_{\|x\|=1} \|A(t)\|$$

The norm of A induced by the Euclidean norm above is equal to the nonnegative square root of the absolute value of the largest eigenvalue of the symmetric matrix $A^T A$. Thus, we define this norm next. The spectral norm of an $m \times n$ matrix A is defined to be

$$\|A\| = \left[\max_{\|x\|=1} x^T A^T A x \right]^{\frac{1}{2}}$$

This will be the matrix norm that is used in the sequel and will be denoted by $\|.\|$.

We now define the concepts of uniform stability and uniform exponential stability. These two concepts involve the boundedness of the solutions of the regressive time varying linear dynamic equation

$$x^{\triangle}(t) = A(t)x(t), \qquad x(t_0) = x_0, \quad t_0 \in \mathbb{T}. \tag{3.1}$$

Definition 1. *The time varying linear dynamic equation (3.1) is uniformly stable if there exists a finite constant $\gamma > 0$ such that for any t_0 and $x(t_0)$, the corresponding solution satisfies*

$$\|x(t)\| \leq \gamma \|x(t_0)\|, \quad t \geq t_0. \tag{3.2}$$

For the next definition, we define a stability property that not only concerns the boundedness of a solution to (3.1), but also the asymptotic characteristics of the solutions as well. If the solution to (3.1) posses the following stability property, then the solution approach zero exponentially as $t \to \infty$ (i.e. the norms of the solutions are bounded above by a decaying exponential function).

Definition 2. *The time varying linear dynamic equation (3.1) is called uniformly exponentially stable if there exists constans $\gamma, \lambda > 0$ with $-\lambda \in \Re^+$ such that for any t_0 and $x(t_0)$, the corresponding solution satisfies*

$$\|x(t)\| \leq \|x(t_0)\| \gamma e_{-\lambda}(t, t_0), \quad t \geq t_0. \tag{3.3}$$

It is obvious by inspection of the previous definitions that we must have $\gamma \geq 1$. By using the word uniform, it is implied that the choice of γ does not depend on the initial time t_0.

The last stability definition given uses a uniformity condition to conclude exponential stability.

Definition 3. *The linear state equation (3.1) is defined to be uniformly asymptotically stable if it is uniformly stable and given any $\delta > 0$, there exists a $T > 0$ so that for any t_0 and $x(t_0)$, the corresponding solution $x(t)$ satisfies*

$$\|x(t)\| \leq \delta \|x(t_0)\|, \quad t \geq t_0 + T. \tag{3.4}$$

It is noted that the time $\mathbb{T}$ that must pass before the norm of the solution satisfies (3.4) and the constant $\delta > 0$ is independent of the initial time t_0.

We now state and prove four theorems, the first three of which characterize uniform stability and uniform exponentially stability in terms of transition matrix for system (3.1). The fourth theorem illustrates the relationship between uniform asymptotic stability and uniform exponential stability.

Theorem 8. *The time varying linear dynamic equation (3.1) is uniformly stable if and only if there exists a $\gamma > 0$ such that*

$\|\Phi_A(t, t_0)\| \leq \gamma$

for all $t \geq t_0$, with $t, t_0 \in \mathbb{T}$.

Theorem 9. *The time varying linear dynamic equation (3.1) is uniformly exponentially stable if and only if there exists* $\gamma, \lambda > 0$ *with* $-\lambda \in \Re^{+}$ *such that*
$\|\Phi_A(t,t_0)\| \leq \gamma e_{-\lambda}(t,t_0)$
for all $t \geq t_0$, *with* $t, t_0 \in \mathbb{T}$.

Theorem 10. *Suppose there exists a constant* α *such that for all* $t \in \mathbb{T}$, $\|A(t)\| \leq \alpha$. *Then the linear state equation (3.1) is called uniformly exponentially stable if and only if there exists a constant* β *such that*

$$\int_{\tau}^{t} \|\Phi_A(t,\sigma(s))\| \bigtriangleup s \leq \beta$$

for all $t \geq t_0$, *with* $t \geq \sigma(\tau)$.

Theorem 11. *The linear state equation (3.1) is uniformly exponentially stable if and only if it is uniformly asymptotically stable.*

3.1. **Perturbation results.** It is also useful to consider state equations that are "close" to another linear state equation that is uniformly stable. In [7, 8], as well as [10], if the stability of system (3.1) has been determined by an appropriate Lyapunov function, then certain conditions on the peturbation matrix $F(t)$ guarantee stability of the perturbed linear system

$$z^{\bigtriangleup}(t) = [A(t) + F(t)]\, z(t). \tag{3.5}$$

DaCunha [4, Theorem 5.1] obtained result about the uniform stability for the perturbed system (3.5) under the condition

$$\int_{\tau}^{\infty} \|F(s)\| \bigtriangleup s \leq \beta.$$

for some $\beta \geq 0$ in the following theorem.

Theorem 12. *Suppose the linear state equations (3.1) is uniformly stable. Then the perturbed linear dynamic equation (3.5) is uniformly stable if there exists some* $\beta \geq 0$ *such that for all* τ

$$\int_{\tau}^{\infty} \|F(s)\| \bigtriangleup s \leq \beta. \tag{3.6}$$

Proof. For any t_0 and $z(t_0) = z_0$, by theorem 2.5 the solution of (3.5) satisfies

$$z(t) = \Phi_A(t,t_0)z_0 + \int_{t_0}^{t} \Phi_A(t,\sigma(s))F(s)Z(s) \bigtriangleup s, \tag{3.7}$$

where $\Phi_A(t,t_0)$ is the transition matrix for system (3.1). By the uniform stability of (3.1), there exists a constant $\gamma > 0$ such that $\|\Phi_A(t,\tau)\| \leq \gamma$, for all $t, \tau \in \mathbb{T}$ with $t \geq \tau$. By taking the norms of both sides of (3.7), we have

$$\|z(t)\| \leq \gamma \|z_0\| + \int_{t_0}^{t} \gamma \|F(s)\| \|Z(s)\| \bigtriangleup s, \qquad t \geq t_0.$$

By Gronwall's Inequality in [3] , a result in [5], and the inequality (3.6), we obtain
$\|z(t)\| \leq \gamma \|z_0\| e_{\gamma\|F(s)\|}(t,t_0)$

$$\leq \gamma \|z_0\| \exp\left(\int_{t_0}^{t} \tfrac{Log(1+\mu(s)\gamma\|F(s)\|)}{\mu(s)} \triangle s\right)$$
$$\leq \gamma \|z_0\| \exp\left(\int_{t_0}^{\infty} \tfrac{Log(1+\mu(s)\gamma\|F(s)\|)}{\mu(s)} \triangle s\right)$$
$$\leq \gamma \|z_0\| \exp\left(\int_{t_0}^{\infty} \gamma \|F(s)\|) \triangle s\right)$$
$$\leq \gamma \|z_0\| e^{\gamma\beta}, \qquad t \geq t_0.$$

Since γ can be used for any t_0 and $z(t_0)$, the state equation (3.5) is uniformly stable. □

4. Main Results

Theorem 13. *Suppose the linear state equations (3.1) is uniformly exponentially stable. Then the perturbed linear dynamic equation (3.5) is uniformly exponentially stable if there exists some $\beta \geq 0$ and $\lambda > 0$ with $-\lambda \in \Re^+$ such that for all $\tau \in \mathbb{T}$*

$$\int_{\tau}^{\infty} \frac{\|F(s)\|}{e_{-\lambda}(\sigma(s), s)} \triangle s \leq \beta. \tag{4.1}$$

Proof. For any t_0 and $z(t_0) = z_0$, by theorem 5 the solution of (3.5) satisfies

$$z(t) = \Phi_A(t, t_0)z_0 + \int_{t_0}^{t} \Phi_A(t, \sigma(s))F(s)Z(s) \triangle s, \tag{4.2}$$

where $\Phi_A(t, t_0)$ is the transition matrix for system (3.1). By the uniform exponentially stability of (3.1), there exists constants $\gamma, \lambda > 0$ with $-\lambda \in \Re^+$ such that $\|\Phi_A(t, \tau)\| \leq \gamma e_{-\lambda}(t, \tau)$, for all $t, \tau \in \mathbb{T}$ with $t \geq \tau$. By taking the norms of both sides of (4.2), we have

$$\|z(t)\| \leq \gamma \|z_0\| e_{-\lambda}(t, t_0) + \int_{t_0}^{t} \gamma e_{-\lambda}(t, \sigma(s)) \|F(s)\| \|Z(s)\| \triangle s, \qquad t \geq t_0.$$

Dividing by $e_{-\lambda}(t, t_0)$ on both sides, we have

$$\tfrac{\|z(t)\|}{e_{-\lambda}(t,t_0)} \leq \gamma \|z_0\| + \gamma \int_{t_0}^{t} \tfrac{e_{-\lambda}(t,\sigma(s))}{e_{-\lambda}(t,t_0)} \|F(s)\| \tfrac{\|Z(s)\|}{e_{-\lambda}(s,t_0)} e_{-\lambda}(s, t_0) \triangle s,$$
$$= \gamma \|z_0\| + \gamma \int_{t_0}^{t} \tfrac{\|F(s)\|}{e_{-\lambda}(\sigma(s),s)} \tfrac{\|Z(s)\|}{e_{-\lambda}(s,t_0)} \triangle s,$$

Letting $u(t) = \frac{\|z(t)\|}{e_{-\lambda}(t,t_0)}$, we have

$$\|u(t)\| \leq \gamma \|z_0\| + \gamma \int_{t_0}^{t} \tfrac{\|F(s)\|}{e_{-\lambda}(\sigma(s),s)} \|u(s)\| \triangle s$$

By Gronwall's Inequality in [3] , a result in [5], and the inequality (4.1), we obtain

$$\leq \gamma \|z_0\| e_{\gamma \frac{\|F(s)\|}{e_{-\lambda}(\sigma(s),s)}}(t,t_0)$$
$$\leq \gamma \|z_0\| \exp\left(\int_{t_0}^{t} \frac{Log(1+\mu(s)\gamma\frac{\|F(s)\|}{e_{-\lambda}(\sigma(s),s)})}{\mu(s)} \triangle s\right)$$
$$\leq \gamma \|z_0\| \exp\left(\int_{t_0}^{\infty} \frac{Log(1+\mu(s)\gamma\frac{\|F(s)\|}{e_{-\lambda}(\sigma(s),s)})}{\mu(s)} \triangle s\right)$$
$$\leq \gamma \|z_0\| \exp\left(\gamma \int_{t_0}^{\infty} \frac{\|F(s)\|}{e_{-\lambda}(\sigma(s),s)} \triangle s\right)$$
$$\leq \gamma \|z_0\| e^{\gamma\beta}, \qquad t \geq t_0.$$

Thus

$$\|z(t)\| \leq \gamma_1 \|z_0\| e_{-\lambda}(t,t_0) \qquad t \geq t_0.$$

where $\gamma_1 = \gamma e^{\gamma\beta}$. Hence the state equation (3.5) is uniformly exponentially stable □

Theorem 14. *Suppose the linear state equations (3.5) is uniformly exponentially stable. Then the perturbed linear dynamic equation (3.5) is uniformly asymptotically stable.*

Proof. By the uniform exponentially stability of (3.5), there exists constants $\gamma, \lambda > 0$ with $-\lambda \in \Re^+$ such that $\|z(t)\| \leq \gamma \|z_0\| e_{-\lambda}(t,t_0)$, for all $t, t_0 \in \mathbb{T}$ with $t \geq t_0$. Clearly, this implies uniform stability. Hence, (3.5) is uniformly stable. Now, given a $\delta > 0$, we choose a sufficiently large positive constant $T \in \mathbb{T}$ so that $t_0 + T \in \mathbb{T}$ and $e_{-\lambda}(t_0+T, t_0) \leq \delta$.

$$\|z(t)\| \leq \gamma \|z_0\| e_{-\lambda}(t,t_0), \qquad t \geq t_0.$$

Then for any t_0 and z_0, and $t \geq t_0 + T$ with $t \in \mathbb{T}$,

$$\leq \gamma e_{-\lambda}(t_0+T,t_0) \|z_0\|$$
$$\leq \delta \|z_0\|, \qquad t \geq t_0 + T.$$

Thus, (3.5) is uniformly asymptotically stable. □

References

[1] R. Agarwal, M. Bohner, D. O'Regan, A. Peterson, Dynamic equations on time scales: a survey, J. Comput. Appl.. Math. 141 (2002), 1-26.

[2] M. Bohner, A. Peterson, Advances in Dynamic Equations on Time Scales, Birkhäuser, Boston, 2003.

[3] M. Bohner, A. Peterson, Dynamic Equations on Time Scales, Birkhäuser, Boston, 2001.

[4] J.J. DaCunha, Stability for time varying linear dynamic systems on time scales, J. Comput. Appl. Math. 176 (2005), 381-410.

[5] T.Gard, J. Hoffaker, Asymptotic behavior of natural growth on time scales, Dynam. Systems Appl. 12 (2003), 131-147.

[6] S. Hilger, Analysis on measure chains- a unified approach to continuous and discrete calculus, Results Math. 18 (1990), 18-56.

[7] R.E. Kalman, J.E. Bertram, Control system analysis and desing via the second method of Lyapunov I: Continuous-time systems, Trans. ASME Ser. D.J. Basic Eng. 82 (1960), 371-393.

[8] R.E. Kalman, J.E. Bertram, Control system analysis and desing via the second method of Lyapunov II: Continuous-time systems, Trans. ASME Ser. D.J. Basic Eng. 82 (1960), 394-400.

[9] W.Kelly, A.Peterson, Difference Equations: An Introduction with Application, Ha rcourt/Academic Press, Burlington, 2001.

[10] W.J. Rung, Linear System Theory, Prentice-Hall, Englewood Cliffs, 1996.

Department of Mathematics, Faculty of Science and Arts, University of Gazi, Beşevler, 06500, Ankara, Turkey
E-mail address: irfany@gazi.edu.tr
E-mail address: adnantuna@gazi.edu.tr, tdast@yahoo.com

JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS,VOL.5,NO.4,347-357,2007,COPYRIGHT 2007 EUDOXUS PRESS,LLC

A Common Fixed Point Theorem of Compatible Maps of Type (α) in Fuzzy Metric Spaces

Servet Kutukcu, Duran Turkoglu, and Cemil Yildiz
Department of Mathematics, Faculty of Science and Arts, Gazi University, 06500 Teknikokullar, Ankara, Turkey
E-mail: skutukcu@yahoo.com, dturkoglu@gazi.edu.tr, cyildiz@gazi.edu.tr

February 20, 2006

Abstract

In this paper we prove a common fixed point theorem for compatible maps of type (α) on fuzzy metric spaces.

Keywords: Fuzzy metric spaces, common fixed point, compatible maps of type (α).
AMS Subject Classifications: 47H10, 54H25

1 Introduction

The notion of fuzzy sets was introduced by Zadeh [30]. Deng [4], Erceg [6], Kaleva and Seikkala [16] and Kramosil and Michalek [19] have introduced the concepts of fuzzy metric spaces in different ways. George and Veeramani [8] modified the concept of fuzzy metric spaces introduced by Kramosil and Michalek [19] in order to get the Hausdorff topology.

Grebiec [9] extended the fixed point theorems of Banach [1] and Edelstein [5] to fuzzy metric spaces in the sense of Kramosil and Michalek [19] whose study is useful in the field of fixed point theorems of contractive type maps. Since then Fang [7] proved some fixed point theorems in fuzzy metric spaces, which improve, generalize and extend some main results of [1,5,10-12,24].

Sessa [25] defined a generalization of commutativity, which called weak commutativity. Further Jungck [14] introduced more generalized commutativity, so called compatibility. Following Grabiec [9], Kramosil and Michalek [19] and Mishra et al. [20] obtained common fixed point theorems for compatible maps and asymptotically commuting maps on fuzzy metric spaces which generalize, extend and fuzzify several fixed point theorems for contractive-type maps on metric spaces and other spaces.

Jungck et al. [15] introduced the concept of compatible maps of type (A) in metric spaces, which is equivalent to the concept of compatible maps under some conditions and proved common fixed point theorems in metric spaces. Cho [2] introduced the notion of compatible maps of type (α) in fuzzy metric spaces.

Many authors have studied the fixed point theory in fuzzy metric spaces. The most interesting references are [7,9,10,11,20,22,26,29].

Recently, Sharma [26] proved a common fixed point theorem for six maps under the condition of compatible maps of type (α) on fuzzy metric spaces.

In this paper, we prove common fixed point theorems for six maps using some conditions in fuzzy metric spaces in the sense of George and Veeramani [8], which turns out to be a

material generalization of the results of Turkoglu et al.[29]. We also give an example to illustrate our main theorem.

2 Preliminaries

Now, we give some definitions.

Definition 2.1. *(Schweizer and Sklar [23]). A binary operation* $* : [0,1]\times[0,1] \to [0,1]$ *is called a continuous t-norm if* $([0,1], *)$ *is an Abelian topological monoid with the unit 1 such that* $a * b \leq c * d$ *whenever* $a \leq c$ *and* $b \leq d$ *for all* $a, b, c, d \in [0,1]$.

Examples of t-norms are $a * b = ab$ and $a * b = \min\{a, b\}$.

Definition 2.2. *(George and Veeramani [8]). The 3-tuble* $(X, M, *)$ *is called a fuzzy metric space (shortly FM-space) if* X *is an arbitrary set,* $*$ *is a continuous t-norm and* M *is a fuzzy set in* $X^2 \times [0,\infty)$ *satisfying the following conditions: for all* $x, y, z \in X$ *and* $t, s > 0$,

(fm-1) $M(x, y, t) > 0$,

(fm-2) $M(x, y, t) = 1$ for all $t > 0$ iff $x = y$,

(fm-3) $M(x, y, t) = M(y, x, t)$,

(fm-4) $M(x, y, t) * M(y, z, s) \leq M(x, z, t + s)$,

(fm-5) $M(x, y, .) : [0,\infty) \to [0,1]$ is continuous.

Note that $M(x, y, t)$ can be thought of as the degree of nearness between x and y with the respect to t. We identify $x = y$ with $M(x, y, t) = 1$ for all $t > 0$ and $M(x, y, t) = 0$ with ∞ and we can find some topological properties and examples of fuzzy metric spaces in [8].

Lemma 2.1. *(Grabiec [9]). For all* $x, y \in X, M(x, y, .)$ *is nondecreasing.*

Definition 2.3. *(Grabiec [9]). Let* $(X, M, *)$ *be an FM-space:*

(1) A sequence $\{x_n\}$ in X is said to be *convergent* to a point x in X (denoted by $lim_{n\to\infty}x_n = x$) if $lim_{n\to\infty}M(x_n, x, t) = 1$ for all $t > 0$.

(2) A sequence $\{x_n\}$ in X is called a *Cauchy sequence* if $lim_{n\to\infty}M(x_{n+p}, x_n, t) = 1$ for all $t > 0$ and $p > 0$.

(3) A FM-space in which every Cauchy sequence is convergent is said to be complete.

Remark 2.1. *Since* $*$ *is continuous, it follows from (fm-4) that the limit of sequence in FM-space is uniquely determined.*

Throughout this paper $(X, M, *)$ will denote the fuzzy metric space with the following condition:

(fm-6) $lim_{t\to\infty}M(x, y, t) = 1$ for all $x, y \in X$ and $t > 0$.

Lemma 2.1. *(Cho [2] and Mishra et al. [20]). Let* $\{y_n\}$ *be a sequence in an FM-space* $(X, M, *)$ *with the condition (fm-6). If there is a number* $k \in (0,1)$ *such that*

$$M(y_{n+2}, y_{n+1}, kt) \geq M(y_{n+1}, y_n, t)$$

for all $t > 0$ *and* $n = 1, 2...$, *then* $\{y_n\}$ *is a Cauchy sequence in* X.

3 Compatible Maps of Type (α)

In this section, we give the concept of compatible maps of type (α) in FM- spaces and some properties of these maps.

DEFINITION 3.1. *(Mishra et al. [20]). Let A and B be maps from an FM-space $(X, M, *)$ into itself. The maps A and B are said to be compatible if*

$$\lim_{n\to\infty} M(ABx_n, BAx_n, t) = 1$$

for all $t > 0$ whenever $\{x_n\}$ is a sequence in X such that

$$\lim_{n\to\infty} Ax_n = \lim_{n\to\infty} Bx_n = z$$

for some $z \in X$.

DEFINITION 3.2. *(Cho [2]). Let A and B be maps from an FM-space $(X, M, *)$ into itself. The maps A and B are said to be compatible of type (α) if*

$$\lim_{n\to\infty} M(ABx_n, BBx_n, t) = 1$$

and

$$\lim_{n\to\infty} M(BAx_n, AAx_n, t) = 1$$

for all $t > 0$ whenever $\{x_n\}$ is a sequence in X such that

$$\lim_{n\to\infty} Ax_n = \lim_{n\to\infty} Bx_n = z$$

for some $z \in X$.

REMARK 3.1. *In [14,15], we can find the equivalent formulations of Definition 4 and 5 and their examples in metric spaces. Such maps are independent of each other and more general than commuting and weakly commuting maps [13,25].*

PROPOSITION 3.1. *(Cho [2]). Let $(X, M, *)$ be an FM-space with $t*t \geq t$ for all $t \in [0, 1]$ and A, B be continuous maps from X into itself. Then A and B are compatible if and only if they are compatible of type (α).*

PROPOSITION 3.2. *(Cho [2]). Let $(X, M, *)$ be an FM-space with $t*t \geq t$ for all $t \in [0, 1]$ and A, B be maps from X into itself. If A and B are compatible of type (α) and $Az = Bz$ for some $z \in X$, then $ABz = BBz = BAz = AAz$.*

PROPOSITION 3.3. *(Cho [2]). Let $(X, M, *)$ be an FM-space with $t*t \geq t$ for all $t \in [0, 1]$ and A, B be compatible maps of type (α) from X into itself. Let $\{x_n\}$ be a sequence in X such that*

$$\lim_{n\to\infty} Ax_n = \lim_{n\to\infty} Bx_n = z$$

for some $z \in X$. Then we have the following:

(i) $lim_{n\to\infty} BAx_n = Az$ if A is continuous at z,

(ii) $ABz = BAz$ and $Az = Bz$ if A and B are continuous at z.

Example. Let the set $X = [0, \infty)$ with the metric d defined by $d(x, y) = |x - y|$ and for each $t > 0$ define $M(x, y, t) = \frac{t}{t+d(x,y)}$ for all $x, y \in X$. Clearly $(X, M, *)$ is a fuzzy metric space where $*$ is defined by $a * b = ab$. Define $A, B : X \to X$ by $Ax = x$ for $x \in [0, 2)$, $Ax = 4$ for $x \in [2, \infty)$, and $Bx = 4 - x$ for $x \in [0, 2)$, $Bx = 4$ for

$x \in [2, \infty)$. Then A and B are discontinuous at $x = 2$. Consider the sequence $\{x_n\}$ in X such that $\lim_n Ax_n = \lim_n Bx_n = z \in X$. By definition of A and B, $z \in [2, \infty)$. Since A and B agree on $[2, \infty)$, we need only consider $z = 2$. Suppose that $x_n \to 2$ and $x_n < 2$ for all n. Then $Ax_n = x_n \to 2$ from the left and $Bx_n = 4 - x_n \to 2$ from the right. Then, since $4 - x_n > 2$ for all n, $ABx_n = 4$ and since $x_n < 2$ for all n, $BAx_n = 4 - x_n \to 2$. Thus $\lim_n M(ABx_n, BAx_n, t) \neq 1$ but $\lim_n M(ABx_n, BBx_n, t) = 1$ and $\lim_n M(BAx_n, AAx_n, t) = 1$ as $x_n \to 2$. Therefore A and B are compatible of type (α) but they are not compatible.

Example. Let the set $X = \mathbb{R}$ with the metric d defined by $d(x, y) = |x - y|$ and for each $t > 0$ define $M(x, y, t) = \frac{t}{t+d(x,y)}$ for all $x, y \in X$. Clearly $(X, M, *)$ is a fuzzy metric space where $*$ is defined by $a * b = ab$. Define $A, B : X \to X$ by $Ax = 1/x^3$ for $x \neq 0$, $Ax = 1$ for $x = 0$, and $Bx = 1/x^2$ for $x \neq 0$, $Bx = 2$ for $x = 0$. Then A and B are discontinuous at $x = 0$. Consider the sequence $\{x_n\}$ in X defined by $x_n = n$, $n = 1, 2, ...$Then we have $\lim_n Ax_n = \lim_n Bx_n = 0$. Further, $\lim_n M(ABx_n, BAx_n, t) = 1$ and $\lim_n M(ABx_n, BBx_n, t) = 0$ and $\lim_n M(BAx_n, AAx_n, t) = 0$. Therefore A and B are compatible but they are not compatible of type (α).

4 Main Results

THEOREM 4.1. *Let $(X, M, *)$ be a complete FM-space with $t * t \geq t$ for all $t \in [0, 1]$ and let A, B, P, S, T and Q be maps from X into itself such that*

(1) $P(X) \subset AB(X), Q(X) \subset ST(X)$,

(2) $AB = BA, ST = TS, PB = BP, QS = SQ, QT = TQ$,

(3) A, B, S and T are continuous,

(4) the pairs (P, AB) and (Q, ST) are compatible of type (α),

(5) there exists a constant $k \in (0, 1)$ such that

$$\begin{aligned} & M^2(Px, Qy, kt) * [M(ABx, Px, kt)M(STy, Qy, kt)] \\ & +aM(STy, Qy, kt)M(ABx, Qy, 2kt) \\ \geq \ & [pM(ABx, Px, t) + qM(ABx, STy, t)]\, M(ABx, Qy, 2kt) \end{aligned}$$

for all x, y in X and $t > 0$ where $0 < p, q < 1$, $0 \leq a < 1$ such that $p + q - a = 1$,

Then A, B, P, S, T and Q have a unique common fixed point in X.

Proof. Let x_0 be an arbitrary point of X. By (1), we can construct a sequence $\{x_n\}$ in X as follows:

$$Px_{2n} = ABx_{2n+1}, Qx_{2n+1} = STx_{2n+2}, n = 0, 1, 2, ...$$

Indeed, such a sequence was first introduced in [27,28].

Now, let $z_n = STx_n$. Then, by (2), we have

$$\begin{aligned} & M^2(Px_{2n+1}, Qx_{2n+2}, kt) \\ & *[M(ABx_{2n+1}, Px_{2n+1}, kt)M(STx_{2n+2}, Qx_{2n+2}, kt)] \\ & +aM(STx_{2n+2}, Qx_{2n+2}, kt)M(ABx_{2n+1}, Qx_{2n+2}, 2kt) \\ \geq \ & [pM(ABx_{2n+1}, Px_{2n+1}, t) + qM(ABx_{2n+1}, STx_{2n+2}, t)] \\ & M(ABx_{2n+1}, Qx_{2n+2}, 2kt) \end{aligned}$$

and

$$\begin{aligned}
& M^2(y_{2n+1}, y_{2n+2}, kt) * [M(y_{2n}, y_{2n+1}, kt)M(y_{2n+1}, y_{2n+2}, kt)] \\
& +aM(y_{2n+1}, y_{2n+2}, kt)M(y_{2n}, y_{2n+2}, 2kt) \\
\geq\ & [pM(y_{2n}, y_{2n+1}, t) + qM(y_{2n}, y_{2n+1}, t)]\, M(y_{2n}, y_{2n+2}, 2kt)
\end{aligned}$$

so

$$\begin{aligned}
& M(y_{2n+1}, y_{2n+2}, kt)\, [M(y_{2n}, y_{2n+1}, kt) * M(y_{2n+1}, y_{2n+2}, kt)] \\
& +aM(y_{2n+1}, y_{2n+2}, kt)M(y_{2n}, y_{2n+2}, 2kt) \\
\geq\ & (p+q)\, M(y_{2n}, y_{2n+1}, t)M(y_{2n}, y_{2n+2}, 2kt)
\end{aligned}$$

hence

$$\begin{aligned}
& M(y_{2n+1}, y_{2n+2}, kt)M(y_{2n}, y_{2n+2}, 2kt) \\
& +aM(y_{2n+1}, y_{2n+2}, kt)M(y_{2n}, y_{2n+2}, 2kt) \\
\geq\ & (p+q)\, M(y_{2n}, y_{2n+1}, t)M(y_{2n}, y_{2n+2}, 2kt)
\end{aligned}$$

and

$$M(y_{2n+1}, y_{2n+2}, kt) + aM(y_{2n+1}, y_{2n+2}, kt) \geq (p+q)\, M(y_{2n}, y_{2n+1}, t)$$

Thus, it follows that

$$M(y_{2n+1}, y_{2n+2}, kt) \geq M(y_{2n}, y_{2n+1}, t)$$

$0 < k < 1$ and for all $t > 0$.

Similarly, we also have

$$M(y_{2n+2}, y_{2n+3}, kt) \geq M(y_{2n+1}, y_{2n+2}, t)$$

$0 < k < 1$ and for all $t > 0$.

In general, for $m = 1, 2, ...$, we have

$$M(y_{m+1}, y_{m+2}, kt) \geq M(y_m, y_{m+1}, t)$$

$0 < k < 1$ and for all $t > 0$. Hence, by Lemma 2.1, $\{y_n\}$ is a Cauchy sequence in X. Since $(X, M, *)$ is complete, it converges to a point z in X. Since $\{Px_{2n}\}$, $\{Qx_{2n+1}\}$, $\{ABx_{2n+1}\}$ and $\{STx_{2n+2}\}$ are subsequences of $\{y_n\}$, they also converge to the point z. Since A and B are continuous and the pair (P, AB) is compatible mappings of type (α), by Proposition 3.3, we have as $n \to \infty$

$$P(AB)x_{2n+1} \to ABz \text{ and } (AB)^2x_{2n+1} \to ABz.$$

Similarly, since S and T are continuous and the pair (Q, ST) is compatible mappings of type (α), by Proposition 3.3, we also have as $n \to \infty$

$$Q(ST)x_{2n+2} \to STz \text{ and } (ST)^2x_{2n+2} \to STz.$$

Now putting $x = (AB)x_{2n+1}$ and $y = x_{2n+2}$ in (5), we have

$$\begin{aligned}
& M^2(P(AB)x_{2n+1}, Qx_{2n+2}, kt) \\
& *[M(AB(AB)x_{2n+1}, P(AB)x_{2n+1}, kt)M(STx_{2n+2}, Qx_{2n+2}, kt)] \\
& +aM(STx_{2n+2}, Qx_{2n+2}, kt)M(AB(AB)x_{2n+1}, Qx_{2n+2}, 2kt) \\
\geq\ & [pM(AB(AB)x_{2n+1}, P(AB)x_{2n+1}, t) + qM(AB(AB)x_{2n+1}, STx_{2n+2}, t)] \\
& M(AB(AB)x_{2n+1}, Qx_{2n+2}, 2kt)
\end{aligned}$$

and

$$\begin{aligned} & M^2(ABz, z, kt) * [M(ABz, ABz, kt)M(z, z, kt)] \\ & +aM(z, z, kt)M(ABz, z, 2kt) \\ \geq\; & [pM(ABz, ABz, t) + qM(ABz, z, t)]\, M(ABz, z, 2kt) \end{aligned}$$

then

$$M^2(ABz, z, kt) + aM(ABz, z, 2kt) \geq [p + qM(ABz, z, t)]\, M(ABz, z, 2kt)$$

and since $M(x, y, .)$ is non-decreasing for all x, y in X, we have

$$\begin{aligned} & M(ABz, z, t)M(ABz, z, 2kt) + aM(ABz, z, 2kt) \\ \geq\; & [p + qM(ABz, z, t)]\, M(ABz, z, 2kt) \end{aligned}$$

hence

$$M(ABz, z, t) \geq 1$$

for all $t > 0$ so $z = ABz$.

Now, we show that $z = Pz$. By putting $x = Px_{2n}$ and $y = x_{2n+1}$ in (5), we have

$$\begin{aligned} & M^2(PPx_{2n}, Qx_{2n+1}, kt) \\ & *[M(AB(Px_{2n}), PPx_{2n}, kt)M(STx_{2n+1}, Qx_{2n+1}, kt)] \\ & +aM(STx_{2n+1}, Qx_{2n+1}, kt)M(AB(Px_{2n}), Qx_{2n+1}, 2kt) \\ \geq\; & [pM(AB(Px_{2n}), PPx_{2n}, t) + qM(AB(Px_{2n}), STx_{2n+1}, t)] \\ & M(AB(Px_{2n}), Qx_{2n+1}, 2kt) \end{aligned}$$

and

$$\begin{aligned} & M^2(Pz, z, kt) * [M(z, Pz, kt)M(z, z, kt)] + aM(z, z, kt)M(z, z, 2kt) \\ \geq\; & [pM(z, Pz, t) + qM(z, z, t)]\, M(z, z, 2kt) \end{aligned}$$

then

$$M^2(Pz, z, kt) + a \geq pM(z, Pz, t) + q.$$

Since $M(x, y, .)$ is non-decreasing and $M(x, y, .) \geq M^2(x, y, .)$ for all x, y in X, we have

$$M(Pz, z, t) + a \geq pM(z, Pz, t) + q$$

hence

$$M(Pz, z, t) \geq 1$$

for all $t > 0$ so $z = Pz$.

By putting $x = Bz$ and $y = x_{2n+1}$ in (5) and using (2), we have

$$\begin{aligned} & M^2(P(Bz), Qx_{2n+1}, kt) \\ & *[M(AB(Bz), (PBz), kt)M(STx_{2n+1}, Qx_{2n+1}, kt)] \\ & +aM(STx_{2n+1}, Qx_{2n+1}, kt)M(AB(Bz), Qx_{2n+1}, 2kt) \\ \geq\; & [pM(AB(Bz), P(Bz), t) + qM(AB(Bz), STx_{2n+1}, t)] \\ & M(AB(Bz), Qx_{2n+1}, 2kt) \end{aligned}$$

and

$$\begin{aligned} & M^2(Bz,z,kt)*[M(Bz,Bz,kt)M(z,z,kt)]+aM(z,z,kt)M(Bz,z,2kt) \\ \geq\ & [pM(Bz,Bz,t)+qM(Bz,z,t)]\,M(Bz,z,2kt) \end{aligned}$$

then

$$M^2(Bz,z,kt)+aM(Bz,z,2kt)\geq [p+qM(Bz,z,t)]\,M(Bz,z,2kt)$$

and since $M(x,y,.)$ is non-decreasing for all x,y in X, we have

$$M(Bz,z,t)M(Bz,z,2kt)+aM(Bz,z,2kt)\geq [p+qM(Bz,z,t)]\,M(Bz,z,2kt)$$

and

$$M(Bz,z,t)+a\geq p+qM(Bz,z,t)$$

hence

$$M(Bz,z,t)\geq 1$$

for all $t>0$ so $z=Bz$. Since $ABz=z$ and $Bz=z$, we have also $Az=z$.

By putting $x=z$ and $y=STx_{2n+2}$ in (5), we have

$$\begin{aligned} & M^2(Pz,Q(ST)x_{2n+2},kt) \\ & *[M(ABz,Pz,kt)M(ST(ST)x_{2n+2},Q(ST)x_{2n+2},kt)] \\ & +aM(ST(ST)x_{2n+2},Q(ST)x_{2n+2},kt)M(ABz,Q(ST)x_{2n+2},2kt) \\ \geq\ & [pM(ABz,Pz,t)+qM(ABz,ST(ST)x_{2n+2},t)] \\ & M(ABz,Q(ST)x_{2n+2},2kt) \end{aligned}$$

and

$$\begin{aligned} & M^2(z,STz,kt)*[M(z,z,kt)M(STz,STz,kt)] \\ & +aM(STz,STz,kt)M(z,STz,2kt) \\ \geq\ & [pM(z,z,t)+qM(z,STz,t)]\,M(z,STz,2kt) \end{aligned}$$

then

$$M^2(z,STz,kt)+aM(z,STz,2kt)\geq [p+qM(z,STz,t)]\,M(z,STz,2kt)$$

so

$$M(z,STz,t)+a\geq p+qM(z,STz,t)$$

hence

$$M(z,STz,t)\geq 1$$

for all $t>0$ so $z=STz$.

To show that $Qz=z$, we take $x=z$ and $y=Qx_{2n+1}$ in (5) and using (2), we have

$$\begin{aligned} & M^2(Pz,QQx_{2n+1},kt) \\ & *[M(ABz,Pz,kt)M(ST(Qx_{2n+1}),QQx_{2n+1},kt)] \\ & +aM(ST(Qx_{2n+1}),QQx_{2n+1},kt)M(ABz,QQx_{2n+1},2kt) \\ \geq\ & [pM(ABz,Pz,t)+qM(ABz,ST(Qx_{2n+1}),t)] \\ & M(ABz,QQx_{2n+1},2kt) \end{aligned}$$

and

$$\begin{aligned} & M^2(z,Qz,kt) * [M(z,z,kt)M(z,Qz,kt)] + aM(z,Qz,kt)M(z,Qz,2kt) \\ \geq \ & [pM(z,z,t) + qM(z,z,t)]\, M(z,Qz,2kt) \end{aligned}$$

then

$$\begin{aligned} & M^2(z,Qz,kt) * M(z,Qz,kt) + aM(z,Qz,kt)M(z,Qz,2kt) \\ \geq \ & (p+q)\, M(z,Qz,2kt) \end{aligned}$$

and

$$\begin{aligned} & M(z,Qz,kt)\, [M(z,Qz,kt) * 1] + aM(z,Qz,kt)M(z,Qz,2kt) \\ \geq \ & (p+q)\, M(z,Qz,2kt) \end{aligned}$$

so

$$M^2(z,Qz,kt) + aM(z,Qz,kt)M(z,Qz,2kt) \geq (p+q)\, M(z,Qz,2kt)$$

and since $M(x,y,.)$ is non-decreasing for all x,y in X, we have

$$M(z,Qz,t) + aM(z,Qz,t) \geq p+q$$

hence

$$M(z,Qz,t) \geq 1$$

for all $t > 0$ so $z = Qz$.

Finally, we show that $Tz = z$. By taking $x = z$ and $y = Tz$ in (5) and using (2), we have

$$\begin{aligned} & M^2(Pz,QTz,kt) * [M(ABz,Pz,kt)M(ST(Tz),QTz,kt)] \\ & +aM(ST(Tz),QTz,kt)M(ABz,QTz,2kt) \\ \geq \ & [pM(ABz,Pz,t) + qM(ABz,ST(Tz),t)] \\ & M(ABz,QTz,2kt) \end{aligned}$$

and

$$\begin{aligned} & M^2(z,Tz,kt) * [M(z,z,kt)M(Tz,Tz,kt)] + aM(Tz,Tz,kt)M(z,Tz,2kt) \\ \geq \ & [pM(z,z,t) + qM(z,Tz,t)]\, M(z,Tz,2kt) \end{aligned}$$

then

$$M^2(z,Tz,kt) + aM(z,Tz,2kt) \geq [p + qM(z,Tz,t)]\, M(z,Tz,2kt)$$

and since $M(x,y,.)$ is non-decreasing for all x,y in X, we have

$$M(z,Tz,t)M(z,Tz,2kt) + aM(z,Tz,2kt) \geq [p + qM(z,Tz,t)]\, M(z,Tz,2kt)$$

so

$$M(z,Tz,t) + a \geq p + qM(z,Tz,t)$$

hence

$$M(z,Tz,t) \geq 1$$

for all $t > 0$ so $z = Tz$. Since $STz = z$ and $z = Tz$, we have also $z = Sz$.

By combining the above results, we have

$$z = Az = Bz = Pz = Sz = Tz = Qz,$$

that is, z is a common fixed point of A, B, P, S, T and Q.

Let v ($v \neq z$) be another common fixed point of A, B, P, S, T and Q. Then on using inequality (5) we have

$$\begin{aligned} & M^2(Pz, Qv, kt) * [M(ABz, Pz, kt)M(STv, Qv, kt)] \\ & +aM(STv, Qv, kt)M(ABz, Qv, 2kt) \\ \geq \; & [pM(ABz, Pz, t) + qM(ABz, STv, t)]\, M(ABz, Qv, 2kt) \end{aligned}$$

so

$$M^2(z, v, kt) + aM(z, v, 2kt) \geq [p + qM(z, v, t)]\, M(z, v, 2kt)$$

and since $M(x, y, .)$ is non-decreasing for all x, y in X, we have

$$M(z, v, t)M(z, v, 2kt) + aM(z, v, 2kt) \geq [p + qM(z, v, t)]\, M(z, v, 2kt)$$

thus, it follows that

$$M(z, v, t) \geq \left(\frac{p-a}{1-q}\right) = 1$$

for all $t > 0$ so $z = v$. Hence A, B, P, S, T and Q have unique common fixed point.

If we put $a = 0$ and $B = T = I_X$ (the identity map on X) in Theorem 4.1, we have the following result in Turkoglu et al. [29];

Corollary 4.1. *Let $(X, M, *)$ be a complete FM-space with $t * t \geq t$ for all $t \in [0, 1]$ and let A, P, S and Q be maps from X into itself such that*

(1) $P(X) \subset S(X), Q(X) \subset S(X)$,

(2) A, S are continuous and $AS = SA$,

(3) the pairs (P, A) and (Q, S) are compatible of type (α),

(4) there exists a constant $k \in (0, 1)$ such that

$$\begin{aligned} & M^2(Px, Qy, kt) * [M(Ax, Px, kt)M(Sy, Qy, kt)] \\ \geq \; & [pM(Ax, Px, t) + qM(Ax, Sy, t)]\, M(Ax, Qy, 2kt) \end{aligned}$$

for all x, y in X and $t > 0$ where $0 < p, q < 1$ such that $p + q = 1$,

Then A, P, S and Q have a unique common fixed point in X.

If we put $P = Q$, $A = S$ and $B = T = I_X$ (the identity map on X) in Theorem 4.1, we have the following result:

Corollary 4.2. *Let $(X, M, *)$ be a complete FM-space with $t * t \geq t$ for all $t \in [0, 1]$ and let P and S be compatible maps of type (α) on X such that $P(X) \subset S(X)$. If S is continuous and there exists a constant $k \in (0, 1)$ such that*

$$\begin{aligned} & M^2(Px, Py, kt) * [M(Sx, Px, kt)M(Sy, Py, kt)] \\ & +aM(Sy, Py, kt)M(Sx, Py, 2kt) \\ \geq \; & [pM(Sx, Px, t) + qM(Sx, Sy, t)]\, M(Sx, Py, 2kt) \end{aligned}$$

for all x, y in X and $t > 0$ where $0 < p, q < 1$, $0 \leq a < 1$ such that $p + q - a = 1$.

Then P and S have a unique common fixed point in X.

If we put $A = B = S = T = I_X$ in Theorem 4.1, we have the following result:

COROLLARY 4.3. *Let $(X, M, *)$ be a complete FM-space with $t * t \geq t$ for all $t \in [0,1]$ and let P and Q be maps from X into itself. If there exists a constant $k \in (0,1)$ such that*

$$\begin{aligned} & M^2(Px, Qy, kt) * [M(x, Px, kt)M(y, Qy, kt)] + aM(y, Qy, kt)M(x, Qy, 2kt) \\ \geq \; & [pM(x, Px, t) + qM(x, y, t)]\, M(x, Qy, 2kt) \end{aligned}$$

for all x, y in X and $t > 0$ where $0 < p, q < 1$, $0 \leq a < 1$ such that $p + q - a = 1$.

Then P and Q have unique common fixed point in X.

Example. Let $X = [0,1]$ with the metric d defined by $d(x,y) = |x - y|$ and for each $t > 0$ define $M(x,y,t) = \frac{t}{t+d(x,y)}$ for all $x, y \in X$. Clearly $(X, M, *)$ is a complete fuzzy metric space where $*$ is defined by $a * b = ab$. Let A, B, P, S, T and Q be maps from X into itself defined as

$$Ax = x, Bx = \frac{x}{2}, Px = \frac{x}{6}, Sx = \frac{x}{5}, Tx = \frac{x}{3}, Qx = 0$$

for all $x \in X$. Then

$$P(X) = \left[0, \frac{1}{6}\right] \subset \left[0, \frac{1}{2}\right] = AB(X) \text{ and } Q(X) = \{0\} \subset \left[0, \frac{1}{15}\right] = ST(X).$$

Clearly $AB = BA, ST = TS, PB = BP, QS = SQ, QT = TQ$ and A, B, S, T are continuous. If we take $k = \frac{1}{2}$ and $t = 1$, we see that the condition (5) of the main Theorem is also satisfied. Moreover, the maps P and AB are compatible of type (α) if $\lim_{n\to\infty} x_n = 0$, where $\{x_n\}$ is a sequence in X such that $\lim_{n\to\infty} Px_n = \lim_{n\to\infty} Sx_n = 0$ for $0 \in X$. Similarly, the maps Q and ST are also compatible of type (α). Thus, all the conditions of main Theorem are satisfied and 0 is the unique common fixed point of A, B, P, S, T and Q.

References

[1] S. Banach, *Theorie les operations lineaires*, Manograie Mathematyezne, Warsaw, Poland, 1932.

[2] Y. J. Cho, *Fixed points in fuzzy metric spaces*, J. Fuzzy Math. 4 (1997) 949-962.

[3] Y. J. Cho, H. K. Pathak, S. M. Kang, J. S. Jung, *Common fixed points of compatible maps of type (β) on fuzzy metric spaces*, Fuzzy Sets and Systems 93 (1998) 99-111.

[4] Z. K. Deng, *Fuzzy pseduo-metric spaces*, J. Math. Anal. Appl. 86 (1982) 74-95.

[5] M. Edelstein, *On fixed and periodic points under contraction mappings*, J. London Math. Soc. 37 (1962) 74-79.

[6] M. A. Erceg, *Metric spaces in fuzzy set theory*, J. Math. Anal. Appl. 69 (1979) 205-230.

[7] J. X. Fang, *On fixed point theorems in fuzzy metric spaces*, Fuzzy Sets and Systems 46 (1992) 107-113.

[8] A. George, P. Veeramani, *On some results in fuzzy metric spaces*, Fuzzy Sets and Systems 64 (1994) 395-399.

[9] M. Grabiec, *Fixed points in fuzzy metric space*, Fuzzy Sets and Systems 27 (1988) 385-389.

[10] O. Hadzic, E. Pap, *Fixed point theory in probabilistic metric spaces*, Kluwer Acad. Publ., 2001.

[11] O. Hadzic, E. Pap, *A fixed point theorem for multivalued mappins in probabilistic metric spaces and an application in fuzzy metric spaces*, Fuzzy Sets and Systems 127 (2002) 333-344.

[12] I. Istratescu, *A fixed point theorem for mappings with a probabilistic contractive iterate*, Rev. Roumaire. Math. Pure Appl. 26 (1981) 431-435.

[13] G. Jungck, *Commuting mappings and fixed points*, Amer. Math. Monthly 83 (1976) 261-263.
[14] G. Jungck, *Compatible mappings and common fixed points*, Internat. J. Math. Math. Sci. 9 (1986) 771-779.
[15] G. Jungck, P. P. Murthy, Y. J. Cho, *Compatible mappings of type (A) and common fixed points*, Math. Japonica 38 (2) (1993) 381-390.
[16] O. Kaleva, S. Seikkala, *On fuzzy metric spaces*, Fuzzy Sets and Systems 12 (1984) 215-229.
[17] E. P. Klement, R. Mesiar, E. Pap, *Triangular norms. Position paper I: Basic analytical and algebraic properties*, Fuzzy Sets and Systems 143 (2004) 5-16.
[18] E. P. Klement, R. Mesiar, E. Pap, *Problems on triangular norms and related operators*, Fuzzy Sets and Systems 145 (2004) 471-479.
[19] O. Kramosil, J. Michalek, *Fuzzy metric and statistical metric spaces*, Kybernetika (Praha) 11 (1975) 326-334.
[20] S. N. Mishra, N. Sharma, S. L. Singh, *Common fixed points of maps on fuzzy metric spaces*, Internat. J. Math. Sci. 17 (1994) 253-258.
[21] H. K. Pathak, Y. J. Cho, S. S. Chang, S. M. Kang, *Compatible mappings of type (P)*, Rev. Res. Univ. Novi Sad., to appear.
[22] B. E. Rhoades, *A comparison of various definitions of contractive mappings*, Trans. Amer. Math. Soc. 226 (1977) 257-290.
[23] B. Schweizer, A. Sklar, *Statistical metric spaces*, Pacific J. Math. 10 (1960) 313-334.
[24] V. M. Sehgal, A. T. Bharucha-Reid, *Fixed point of contraction mapping on PM spaces*, Math. Systems Theory 6 (1972) 97-100.
[25] S. Sessa, *On weak commutativity condition of mappings in fixed point considerations*, Publ. Inst. Math. Beagrad 32 (46) (1982) 149-153.
[26] S. Sharma, *Common Fixed point theorems in fuzzy metric spaces*, Fuzzy Sets and Systems 127 (2002) 345-352.
[27] S. L. Singh, S. Kasahara, *On some recent results on common fixed points*, Indian J. Pure Appl. Math. 13 (1982) 757-761.
[28] S. L. Singh, B. Ram, *Common fixed points of commuting mappings in 2-metric spaces*, Math. Sem. Notes Kobe Univ. 10 (1982) 197-208.
[29] D. Turkoglu, S. Kutukcu, C. Yildiz, *Common fixed points of compatible maps of type (α) on fuzzy metric spaces*, Int. J. Appl. Math., in press.
[30] L. A. Zadeh, *Fuzzy sets*, Inform. and Control 8 (1965) 338-353.

Instructions to Contributors
Journal of Concrete and Applicable Mathematics

A quartely international publication of Eudoxus Press, LLC, of TN.

Editor in Chief: George Anastassiou

Department of Mathematical Sciences
University of Memphis
Memphis, TN 38152-3240, U.S.A.

1. Manuscripts hard copies in triplicate, and in English, should be submitted to the Editor-in-Chief:

Prof.George A. Anastassiou
Department of Mathematical Sciences
The University of Memphis
Memphis,TN 38152, USA.
Tel.001. 901.678.3144
e-mail:ganastss@memphis.edu.

Authors may want to recommend an associate editor the most related to the submission to possibly handle it.
Also authors may want to submit a list of six possible referees, to be used in case we cannot find related referees by ourselves.
2. Manuscripts should be typed using any of TEX,LaTEX,AMS-TEX,or AMS-LaTEX and according to EUDOXUS PRESS, LLC. LATEX STYLE FILE. (VISIT www.msci.memphis.edu/~ganastss/jcaam / to save a copy of the style file.)They should be carefully prepared in all respects. Submitted copies should be brightly printed (not dot-matrix), double spaced, in ten point type size, on one side high quality paper 8(1/2)x11 inch. Manuscripts should have generous margins on all sides and should not exceed 24 pages.
3. Submission is a representation that the manuscript has not been published previously in this or any other similar form and is not currently under consideration for publication elsewhere. A statement transferring from the authors(or their employers,if they hold the copyright) to Eudoxus Press, LLC, will be required before the manuscript can be accepted for publication.The Editor-in-Chief will supply the necessary forms for this transfer.Such a written transfer of copyright,which previously was assumed to be implicit in the act of submitting a manuscript,is necessary under the U.S.Copyright Law in order for the publisher to carry through the dissemination of research results and reviews as widely and effective as possible.
4. The paper starts with the title of the article, author's name(s) (no titles or degrees), author's affiliation(s) and e-mail addresses. The affiliation should comprise the department, institution (usually university or company), city, state (and/or nation) and mail code.
The following items, 5 and 6, should be on page no. 1 of the paper.

5. An abstract is to be provided, preferably no longer than 150 words.
6. A list of 5 key words is to be provided directly below the abstract. Key words should express the precise content of the manuscript, as they are used for indexing purposes. The main body of the paper should begin on page no. 1, if possible.
7. All sections should be numbered with Arabic numerals (such as: 1. INTRODUCTION) .
Subsections should be identified with section and subsection numbers (such as 6.1. Second-Value Subheading).
If applicable, an independent single-number system (one for each category) should be used to label all theorems, lemmas, propositions, corrolaries, definitions, remarks, examples, etc. The label (such as Lemma 7) should be typed with paragraph indentation, followed by a period and the lemma itself.
8. Mathematical notation must be typeset. Equations should be numbered consecutively with Arabic numerals in parentheses placed flush right,and should be thusly referred to in the text [such as Eqs.(2) and (5)]. The running title must be placed at the top of even numbered pages and the first author's name, et al., must be placed at the top of the odd numbed pages.
9. Illustrations (photographs, drawings, diagrams, and charts) are to be numbered in one consecutive series of Arabic numerals. The captions for illustrations should be typed double space. All illustrations, charts, tables, etc., must be embedded in the body of the manuscript in proper, final, print position. In particular, manuscript, source, and PDF file version must be at camera ready stage for publication or they cannot be considered.
Tables are to be numbered (with Roman numerals) and referred to by number in the text. Center the title above the table, and type explanatory footnotes (indicated by superscript lowercase letters) below the table.
10. List references alphabetically at the end of the paper and number them consecutively. Each must be cited in the text by the appropriate Arabic numeral in square brackets on the baseline.
References should include (in the following order):
initials of first and middle name, last name of author(s)
title of article,
name of publication, volume number, inclusive pages, and year of publication.
Authors should follow these examples:

<u>Journal Article</u>

1. H.H.Gonska,Degree of simultaneous approximation of bivariate functions by Gordon operators, (journal name in italics) *J. Approx. Theory*, 62,170-191(1990).

<u>Book</u>

2. G.G.Lorentz, (title of book in italics) *Bernstein Polynomials* (2nd ed.), Chelsea,New York,1986.

<u>Contribution to a Book</u>

3. M.K.Khan, Approximation properties of beta operators,in(title of book in italics) *Progress in Approximation Theory* (P.Nevai and A.Pinkus,eds.), Academic Press, New York,1991,pp.483-495.

11. All acknowledgements (including those for a grant and financial support) should occur in one paragraph that directly precedes the References section.
12. Footnotes should be avoided. When their use is absolutely necessary, footnotes should be numbered consecutively using Arabic numerals and should be typed at the bottom of the page to which they refer. Place a line above the footnote, so that it is set

off from the text. Use the appropriate superscript numeral for citation in the text.
13. After each revision is made please again submit three hard copies of the revised manuscript, including in the final one. And after a manuscript has been accepted for publication and with all revisions incorporated, manuscripts, including the TEX/LaTex source file and the PDF file, are to be submitted to the Editor's Office on a personal-computer disk, 3.5 inch size. Label the disk with clearly written identifying information and properly ship, such as:

Your name, title of article, kind of computer used, kind of software and version number, disk format and files names of article, as well as abbreviated journal name.

Package the disk in a disk mailer or protective cardboard. Make sure contents of disks are identical with the ones of final hard copies submitted!

Note: The Editor's Office cannot accept the disk without the accompanying matching hard copies of manuscript. No e-mail final submissions are allowed! The disk submission must be used.

14. Effective 1 Nov. 2005 the journal's page charges are $10.00 per PDF file page. Upon acceptance of the paper an invoice will be sent to the contact author. The fee payment will be due one month from the invoice date. The article will proceed to publication only after the fee is paid. The charges are to be sent, by money order or certified check, in US dollars, payable to Eudoxus Press, LLC, to the address shown in the Scope and
Prices section.

No galleys will be sent and the contact author will receive one(1) electronic copy of the journal issue in which the article appears.

15. This journal will consider for publication only papers that contain proofs for their listed results.

TABLE OF CONTENTS,JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS,VOL.5,NO.4,2007